New methods in nonequilibrium gases and fluids[@]

Giovanni Gallavotti*

* Fisica, Università di Roma, P.le Moro 2, 00185 Roma

ABSTRACT. Kinematical and dynamical properties of chaotic systems are reviewed and a few applications are described.

§*1 Kinematics.*

Gases and fluids are described mathematically by dynamical systems. This means that the system state is a point in a *phase space* $\mathcal{C}$ on which the motion is given by a map S.

Typically $\mathcal{C}$ is a finite dimensional smooth finite (*i.e.* compact) manifold and S is a local diffeomorphism of $\mathcal{C}$. But sometimes $\mathcal{C}$ and/or S have singularities (think for instance of hard sphere systems) located on a closed set N of zero volume: in such cases the points of $\mathcal{C}/N$ are the *regular points* where S is smooth, of class C^∞, with non vanishing jacobian determinant and N will always contain the boundary of $\mathcal{C}$, if any.

The transformation S can be *invertible*: this means that there is another map S^{-1} with a singularity set N' (with 0 volume) such that if $x \notin N$ and $Sx \notin N'$ then $S^{-1}Sx = x$, and if $x \notin N'$ and $S^{-1}x \notin N'$ then $SS^{-1}x = x$.

Unless explicitly stated we always suppose the maps we consider to be invertible and without singularities.

If there is a smooth map i of $\mathcal{C}$ into itself such that $i^2 = 1$, and $iS = S^{-1}i$ we say thgat the system is time reversal symmetric. We can always suppose that i is an isometry (possibly changing the distance $d(x,y)$ to $\tilde{d}(x,y) = d(x,y) + d(ix,iy)$).

The main question is: *given* $(\mathcal{C}, S)$ *and picking* $x \in \mathcal{C}$ *randomly with a distribution* μ_0 *proportional to the volume* dx: $\mu_0(dx) = \rho_0(x)\,dx$ *(for some* $\rho_0 \in L_1(\mathcal{C}, dx)$*), which is the asymptotic behaviour of the motion starting at* x*?*

This means asking what can one say about the averages:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{j=0}^{T-1} f(S^{\pm j}x) = \langle f \rangle_\pm \tag{1.1}$$

In the case of a non invertible transformation one chooses, of course, only the sign $+$ in (1.1).

We shall always suppose that the averages exist for μ_0–almost all x's: this assumpiton is called 0-*th law*. The necessity of the clause "μ_0–almost everywhere" is because in

[@] This paper is dedicated to the memory of Giovanni Paladin, in grief for the untimely end of his life, 29 june 1996.

essentially all cases there are obvious exceptions. Typically each periodic orbit is an exception and usually periodic orbits form a dense set in phase space (the cases with no periodic orbits, as the quasi periodic motions, must be regarded as exceptional).

The independence on x of the r.h.s. of (1.1), *i.e.* the uniqueness of the averages, is not a very strong assumption. Usually it fails because $\mathcal{C}$ breaks into open parts $U_1, U_2, \ldots$ whose union has full volume and in each of them the asymptotic behavior is unique, *i.e.* for each j one has x–independence almost everywhere of the average μ_0–almost everywhere in U_j. Thus the system would, essentially, break down into several systems each of which verifies the 0–th law.

The averages $\langle f\rangle_\pm$ can be written:

$$\langle f\rangle_\pm = \int_{\mathcal{C}} f(y)\mu_\pm(dy) \tag{1.2}$$

where $\mu_\pm$ is a *probability distribution* that will be called the *forward statistics* (+) and the *backward statistics* (−) of μ_0 or the statistics of the motions with respect to initial data choices at random with respect to the volume in phase space.

Are there systems $(\mathcal{C}, S)$ for which μ can be determined?

(1) *Hamiltonian systems*: in this case the phase space is the space of the configurations and momenta x of given energy that verify some *"timing property"*, for instance x lies on a certain surface signaling some interesting event, *e.g.* a collision. If $t(x)$ is the time interval between the "event" x and the next one of the same type the map S is expresssed in terms of the solution flow $t \to S_t$ of the equations of motion by $S = S_{t(x)}$. The Liouville theorem gives us a probability distribution μ_0 which is proportional to the volume on $\mathcal{C}$ and it is *invariant*: $\mu_0(E) = \mu_0(S^{-1}E)$ for all sets E for which it makes sense to measure the volume ("volume measurable sets").

The *ergodic hypothesis*, when reasonable, tells us that μ_0 is the statistics of itself: $\mu_+ = \mu_0$, [1].

(2) *Transitive Anosov systems*: in this case the phase space is a smooth manifold (say C^∞) and through each point x pass a smooth stable manifold W^s_x and a smooth unstable manifold W^u_x such that their tangent planes T^s_x, T^u_x at x:

(i) are S–covariant (*i.e.* $\partial S\, T^u_x = T^u_{Sx}, \partial S T^s_x = T^s_{Sx}$), vary continuously with x, in the sense that T^s_x, T^u_x are Hölder continuous as functions of x,[1] and the full tangent plane is the direct sum of T^s_x, T^u_x.

(ii) for all $n \ge 0$:

$$\begin{array}{lll} (a) & ||\partial\, S^n\, \underline{v}\,||_{S^n x} \le Ce^{-n\lambda}||\,\underline{v}\,||_x & \underline{v} \in T^s_x \\ (b) & ||\partial\, S^{-n}\, \underline{v}\,||_{S^n x} \le Ce^{-n\lambda}||\,\underline{v}\,||_x & \underline{v} \in T^u_x \end{array} \tag{1.3}$$

(iii) There is a periodic point with stable and unstable manifolds dense in $\mathcal{C}$.

It then follows that all points have dense stable and unstable manifolds and that the periodic points are dense.

For such systems the 0–th law holds as a mathematical theorem, [2], [3], p. 757.

[1] One might prefer to require real smoothness, *e.g.* C^p with $1 \le p \le \infty$: but this would be too much for rather trivial reasons. On the other hand Hölder continuity might be equivalent to simple C^0–continuity as in the case of Anosov systems, see [4], [3].

(3) *Transitive Axiom A attractors*: Suppose that $(\mathcal{C}, S)$ is a smooth system and $\mathcal{S}$ is a smooth invariant surface which is attracting in the sense that the distance of $S^n x$ from $\mathcal{S}$ tends to 0 exponentially: $d(S^n, \mathcal{S}) \le Ce^{-\lambda n} d(x, \mathcal{S})$ for all x in the vicinity of $\mathcal{S}$ (with $C, \lambda > 0$ suitable constants). If $(\mathcal{S}, S)$ is a transitive Anosov system then this is an example of a *transitive Axiom A attractor*. More generally an Axiom A attractor needs not be a smooth manifold; it can be a non smooth closed attracting set $\mathcal{A}$ provided:

(a) at each point $x \in \mathcal{A}$ the tangent space to $\mathcal{C}$ can be decomposed into two planes T_x^s, T_x^u so that (i), (ii) above hold, and (iii) is replaced by the requirement that on $\mathcal{A}$ there is a periodic point with dense stable and unstable manifolds.

For such systems the 0–th law holds and the statistics $\mu_\pm$ are called *SRB distributions*.
In general we call *attracting set* a closed invariant set $\mathcal{A}$ such that $d(S^n x, \mathcal{A}) \xrightarrow[n\to+\infty]{} 0$ for all x sufficiently close to $\mathcal{A}$, and furthermore no subset of $\mathcal{A}$ enjoys the same property. A *repeller* is defined analogously by using S^{-1} instead of S.
Note that if $\mathcal{A}$ is an Axiom A attractor then the planes T_x^u are "tangent to the attractor" in the sense that they are everywhere tangent to a smooth invariant surface W_x^u contained in $\mathcal{A}$. But the plane T_x^s are not tangent to $\mathcal{A}$: in fact part of them sticks out of $\mathcal{A}$ (describing the fall on the attractor of the nearby points) and part of them lies in some sense on the attractor and describes motions that stay on the attractor and are *homoclinic* as $n \to +\infty$, *i.e.* approach each other as $n \to +\infty$.
In general through every point in an Axiom A attractor pass two smooth manifolds W_x^s, W_s^u which are tangent to T_x^s, T_x^u.

(4) *Axiom A systems*: A notion more general than the previous one can be given by introducing the *non wandering points*: they are the points x such that any of their neighborhoods returns infinitely often near x. This means that for any pair of neighborhoods U, V of x there are infinitely many $n > 0$ such that $S^n V \cap U \neq \emptyset$, [3] p.749.
A smooth system $(\mathcal{C}, S)$ verifies Axiom A if the set Ω of nonwandering points is hyperbolic, *i.e.* it verifies property (i), (ii) of the previous example (2) and (iii) is replaced by the requirement that the periodic points are dense, [3] p.777, [5] p. 154.
If an Axiom A system has an attractive set then it is called an "axiom A attractor" (not necessarily transitive, however see below). In general through every non wandering point in an Axiom A system pass two smooth manifolds W_x^s, W_s^u which are tangent to T_x^s, T_x^u.
The nonwandering set of an Axiom A system splits into a finite union of invariant sets, called *basic sets*, $\{\Omega_j\}$ on each of which there is a dense orbit ("topological transitivity", not to be confused with the above notion of transitivity). And each of such sets splits into a finite union of sets $\{B_{jk}\}$ each of which is invariant for an appropriate *iterate* S^{p_j} which acts transitively on it (in the sense of the previous examples: namely the stable and unstable manifolds of a periodic point are dense). Thus if Ω_j is an attractor each B_{jk} is an Axiom A transitive attractor for S^{p_j}).
The 0–th law holds for the attractors of an Axiom A system.

(5) *Axiom C systems*: The Axiom A and the zeroth law properties could be reasonably taken as properties characterizing models for globally "chaotic" or "globally hyperbolic" systems. But in Axiom A systems there are no relations between the different basic sets. And therefore other more global kinematic notions have been considered in the literature: we shall not dwell here on the Axiom B notion and we deal directly with the Axiom C notion that is particularly appropriate for invertible maps enjoying a time reversal symmetry. In fact the problems that we pose in the next sections suggest that

the appropriate notion for "globally hyperbolic" or "globally chaotic" dynamical systems is somewhat stronger than that of Axiom A or B.

Axiom C systems are Axiom A systems enjoying further properties. To describe them we introduce the notion of distance of a point x to the basic sets $\{\Omega_i\}$ of an Axiom A system as:

$$\delta(x) = \min\{\min_i \frac{d_{\Omega_i}(x)}{d_0}, \inf_{j,\, -\infty<n<+\infty} \frac{d_{\Omega_j}(S^n x)}{d_0}\} \tag{1.4}$$

where d_0 is the diameter of the phase space $\mathcal{C}$, $d_{\Omega_i}(x)$ is the distance of the point x from the basic set Ω_i, the minimum over i runs over the attracting or repelling basic sets and the minimum over j runs over the other basic sets (if any).

We can then define the *Axiom C* systems as:

A smooth dynamical system $(\mathcal{C}, S)$ verifies Axiom C if it verifies Axiom A and:
(1) among the basic sets there are a unique attracting and a unique repelling basic sets, denoted Ω_+, Ω_- respectively, with (open) full volume dense basins of attraction. We call $\Omega_\pm$ the *poles* of the system (future or attracting and past or repelling poles, respectively).
(2) for every $x \in \mathcal{C}$ the tangent space T_x admits a Hölder–continuous decomposition as a direct sum of three subspaces T_x^u, T_x^s, T_x^m such that:

a) $\partial S\, T_x^\alpha \quad = T_{Sx}^\alpha \qquad\qquad \alpha = u, s, m$
b) $|\partial S^n w| \quad \le C e^{-\lambda n} |w|, \qquad w \in T_x^s,\ n \ge 0$
c) $|\partial S^{-n} w| \quad \le C e^{-\lambda n} |w|, \qquad w \in T_x^u,\ n \ge 0$
d) $|\partial S^n w| \quad \le C \delta(x)^{-1} e^{-\lambda |n|} |w|, \qquad w \in T_x^m,\ \forall n$

where the dimensions of T_x^u, T_x^s, T_x^m are > 0 and $\delta(x)$ is defined in (1.4); here ∂S^n is the jacobian matrix of S^n.
(3) if x is on the attracting pole Ω_+ then $T_x^s \oplus T_x^m$ is tangent to the stable manifold in x; viceversa if x is on the repelling pole Ω_- then $T_x^u \oplus T_x^m$ is tangent to the unstable manifold in x.

Although T_x^u and T_x^s are not uniquely determined by the above definition the planes $T_x^s \oplus T_x^m$ and $T_x^u \oplus T_x^m$ are uniquely determined for $x \in \Omega_+$ and, respectively, $x \in \Omega_-$.

The above notion of Axiom C is in [6]. It is clear that an Axiom C system is necessarily also an Axiom A system verifying the 0–th law.

If Ω_+ and Ω_- are the two poles of the system the stable manifold of a periodic point $p \in \Omega_+$ and the unstable manifold of a periodic point $q \in \Omega_-$ not only have a point of transversal intersection (this would be the property characterizing Axiom B systems among the Axiom A ones with only two basic sets) but they intersect transversally *all the way* on a manifold connecting Ω_+ to Ω_-; the unstable manifold of a point in Ω_- will accumulate on Ω_+ *without winding around it*, [6].

In fact one can *"attach"* to W_p^s, $p \in \Omega_+$, points on Ω_- as follows: we say that a point $z \in \Omega_-$ is *attached* to W_p^s if it is an accumulation point for W_p^s and there is a closed curve *with finite length* linking a point $z_0 \in W_p^s$ to z and *entirely lying on* W_p^s, with the exeption of the endpoint z. A drawing helps understanding this simple geometrical construction, slightly unusual because of the density of W_p^s on Ω_-,.

We call $\overline{W}_p^s$ the set of the points *either on* W_p^s *or just attached to* W_p^s on the system basic sets (the set $\overline{W}_p^s$ should not be confused with the closure clos(W_p^s), which is the whole space, see [3], p. 783). If a system verifies Axiom C the set $\overline{W}_p^s$ intersects Ω_- on a stable manifold, by 2) in the above definition.

The definition of $\overline{W}_q^u$, $q \in \Omega_-$, is set symmetrically by exchanging Ω_+ with Ω_-.

Furthermore if a system verifies Axiom C and $p \in \Omega_+$, $q \in \Omega_-$ are two periodic points, on the attracting and on the repelling pole of the system respectively, then $\overline{W}^s_p$ and $\overline{W}^u_q$ have a dense set of points in Ω_- and Ω_+, respectively.

Note that $\overline{W}^s_p \cap \overline{W}^u_q$ is dense in $\mathcal{C}$ as well as in Ω_+ and Ω_-. This follows from the density of W^s_p and W^u_q on Ω_+ and Ω_- respectively and from the continuity of T^m_x. Furthermore if $z \in \Omega_+$ is such that $z \in \overline{W}^s_p \cap \overline{W}^u_q$ then the surface $\overline{W}^s_p \cap \overline{W}^u_q$ intersects Ω_- in a unique point $\tilde{z} \equiv \tilde{\imath} z$ which can be reached by the shortest smooth path on $\overline{W}^s_p \cap \overline{W}^u_q$ linking z to C_- (the path is on the surface obtained as the envelope of the tangent planes T^m, but it is in general not unique even if T_m has dimension 1, see the example in Appendix A1 below).

The map $\tilde{\imath}$ commutes with S, squares to the identity, maps Ω_+ to Ω_- and viceversa and will play a key role in the following analysis.

In [6] it is conjectured that the Axiom C systems are Ω–stable in the sense of Smale, [3] p. 749, *i.e.* small perturbations of Axiom C systems are still Axiom C systems.

§*2 Dynamics and chaotic hypothesis*

We begin with a few examples of systems that we would like to study.

Example 1: We shall consider the problem of electrical conduction in a crystal via the classical model representing the crystal as a periodic array of circular obstacles among which free charges, also modeled by hard spheres, move undergoing elastic collisions with the obstacles and between themselves.

An electric field acts upon the moving particles establishing a current. We suppose that the obstacles are such that there is no straight line that avoids them. Nevertheless clearly the electric field will continue to work so that the electric current will grow unbounded and therefore the system will show infinite conductivity.

The reason why electrical conductivity is finite in physical systems is simply that there are *dissipative effects*. The simplest theories of conductivity, like Drude's theory, write directly a dissipative equation for the motion of the N moving particles contained in a crystal with cell of size a:

$$\dot{\underline{q}}_j = \frac{\underline{p}_j}{m}, \qquad \dot{\underline{p}}_j = \underline{F}_j + E\,\underline{i} - \nu\,\underline{p}_j \tag{2.1}$$

where $\nu = 2\ell v^{-1}$ if ℓ is the mean free path (equal to $\ell = (4\pi\rho d^2)^{-1}$ if d is the radius of the particles and $\rho = N/a^3$ their density) and v is the average velocity, F_j is the (impulsive) force acting on the j–th particle. This special choice of ν corresponds to a dissipation proportional to the speed: no special physical meaning should be attached to this choice which is here used only as an example.

Periodic boundary conditions are imposed at the cell boundary in the direction parallel to the field and reflecting conditions are imposed in the direction orthogonal to the field (*semiperiodic boundary conditions*), to fix the ideas.

The above are not the only equations that we can use for the conduction problem: in fact the coefficient ν has an empirical nature and it is supposed to model a thermostat action that forbids the system to reach arbitrarily high current.

At the same level of rigor one could equally well maintain that the thermostat action is that of keeping the total energy of the moving particles constant; hence one can think of imposing the anholonomic constraint $\sum_j \underline{p}_j^2 = const$ and the constant should be $\mathcal{E} = N(\frac{3}{2}k_B T + \frac{1}{2}m\overline{v}^2)$ if T is the absolute temperature, k_B the Boltzmann constant and $\overline{v}$ is the average drift velocity.

Gauss asked himself what would be the way of imposing a nonholonomous constraint affecting in a "minimal fashion" the dynamics and formulated the well known extension of D'Alembert's principle called the *least constraint* principle. If the constraint of constant energy is imposed on the above system, by Gauss' principle, the resulting equations take the form (2.1) with ν replaced by a multiplier α given by:

$$\alpha(\underline{p}) = \frac{E\,\underline{i}\cdot\sum_j \underline{p}_j}{\sum_j \underline{p}_j^2} \tag{2.2}$$

A brief computations shows that indeed if in (2.1) one substitutes $\nu \to \alpha$ the solutions of the new equation (2.1) keep a constant total (kinetic) energy. This model was considered for instance in [7].

Of course the value of the constant given to the energy cannot be arbitrary. Suppose that the dissipative effects modeled by ν are such that when the external field is E then the system relaxes to a stationary state in which the average energy is $\mathcal{E}$. Then if we want that the model (2.1) or the same model with ν replaced by the α of (2.2) to be equivalent we should fix the energy in the second model to be exactly $\mathcal{E}$. For large systems the two models should be equivalent, [8].

Example 2: A more physical model for electric conduction is:

$$\dot{\underline{q}}_j = \underline{p}_j/m, \qquad \dot{\underline{p}}_j = \underline{F}_j - \beta\,\underline{p}_j + m\,\dot{\underline{B}}_j(t) + E\,\underline{i} \tag{2.3}$$

where $\underline{B}_j$ is a brownian motion. This is a *Langevin equation* in which β is a friction coefficient (*e.g.* ν above) and the dispersion δ of each component of $\dot{\underline{B}}_j$ is related to the temperature (by $\frac{1}{2}\frac{m\delta^2}{2\beta} = \frac{1}{2}k_B T$, see [9], p. 55).

The above equation is clearly phenomenological and one should think that the same physical system may be described in a different way, for instance by the system of equations:

$$\begin{cases} \dot{\underline{q}}_j = \underline{p}_j/m \\ \dot{\underline{p}}_j = \underline{F}_j - \alpha\,\underline{p}_j + m\,\dot{\underline{B}}_j(t) + E\,\underline{i} \end{cases}, \qquad \begin{cases} \dot{\underline{b}}_{jk} = \underline{B}_{jk} \\ \dot{\underline{B}}_{jk} = -\underline{\partial}_{\,\underline{b}_{jk}} V(\{\underline{b}_{j'k'}\}) \end{cases}$$
$$\underline{B}_j \stackrel{def}{=} \sum_k \underline{B}_{jk} \tag{2.4}$$

where the multiplier α is determined by imposing that the total kinetic energy $\mathcal{E} = \sum_j \frac{1}{2m}\underline{p}_j^2$ of the charged particles stays constant:

$$\alpha = \frac{\sum_j (E\,\underline{i} - \dot{\underline{B}}_j)\cdot\underline{p}_j}{\sum_j \underline{p}_j^2} \tag{2.5}$$

and the $(\underline{b}, \underline{B})$ coordinates describe a hamiltonian system with many degrees of freedom (*i.e.* the label k takes many values, *e.g.* $k \gg N$) which acts as a *heat reservoir* on the system of particles.

Again the equivalence between the two equations can hold only if the value $\mathcal{E}$ of the energy in (2.4) is fixed equal to the average value of the energy of (2.3) and, at the same time, the motion of the (hamiltonian) dynamical system for the evolution of the variables $\underline{b}, \underline{B}$ is so chaotic that the variables $\underline{B}_j$ can be reagrded as a gaussian process whose covariance is the same, or close, to that of the homomimous stochastic process in (2.3), [8].

Example 3 As third and last example we shall focus on fluid mechanics problems considering a fluid that:
(1) is enclosed in a periodic box Ω with side L, possibly with a few disks ("obstacles") removed so that no infinite straight path can be found in Ω that avoids the obstacles,
(2) is incompressible with density ρ.
I shall consider two distinct evolution equations for this fluid, both of dissipative nature.

$$\begin{aligned}
&\dot{\underline{u}} + \underset{\sim}{u} \cdot \underset{\sim}{\partial}\, \underline{u} = -\frac{1}{\rho}\underline{\partial} p + \underline{g} + \nu\Delta\underline{u}, \qquad \underline{\partial}\cdot\underline{u} = 0 \qquad \text{NS}\\
&\dot{\underline{u}} + \underset{\sim}{u} \cdot \underset{\sim}{\partial}\, \underline{u} = -\frac{1}{\rho}\underline{\partial} p + \underline{g} + \beta\Delta\underline{u}, \qquad \underline{\partial}\cdot\underline{u} = 0 \qquad \text{GNS}
\end{aligned} \tag{2.6}$$

In the case Ω *contains obstacles a "no friction" boundary condition will be imposed on* $\partial\Omega$, *i.e.* $\underline{u}\cdot\underline{n} = 0$ if $\underline{n}$ is the normal to $\partial\Omega$. The first equation is the well known Navier Stokes equation with ν being the *viscosity*.
The second equation, introduced in [8] and called the gaussian Navier Stokes equation or GNS equation, has a multiplier β defined so that the total vorticity $\eta L^3 = \rho\int \underline{\omega}^2\, dx$, with $\underline{\omega} = \underline{\partial}\wedge\underline{u}$ being the *vorticity*, is a constant of motion; this means that:

$$\beta(\underline{u}) = \frac{\int\big(\underline{\partial}\wedge\underline{g}\cdot\underline{\omega} + \underline{\omega}\cdot(\underset{\sim}{\omega}\cdot\underset{\sim}{\partial}\,\underline{u})\big)\,d\underline{x}}{\int(\underline{\partial}\wedge\underline{\omega})^2\,d\underline{x}} \stackrel{def}{=} \beta_e + \beta_i \tag{2.7}$$

The above equations were studied in [10]. We shall consider only the *truncated equations* with momentum cut off K. In this way the existence and uniquenes problems are completely avoided.
The truncation is performed on a suitable orthonormal basis for the *divergenceless* fields in Ω: given the boundary conditions we consider it natural to use the basis generated by the *minimax* principle applied to the Dirichlet quadratic form $\int_\Omega(\underset{\sim}{\partial}\;\underline{u})^2 d\underline{x}$ defined on the space of the $C^\infty(\Omega)$ divergenceless fields $\underline{u}$ with $\underline{u}\cdot\underline{n} = 0$ on $\partial\Omega$. The basis fields $\underline{u}_j$ will verify: $\Delta\underline{u}_j = -E_j\underline{u}_j + \underline{\partial}_j\mu$, for a suitable multiplier μ_j, with $\underline{u}_j, \mu_j \in C^\infty$ and E_j are eigenvalues). Thus truncating at momentum K means setting identically 0 the components on $\underline{u}_j$ with $E_j > K^2$.
For instance in the case of *no obstacles* let the $\underline{u} = \sum_{\underline{k}\neq\underline{0}}\underline{\gamma}_{\underline{k}} e^{i\underline{k}\cdot\underline{x}}$ be the velocity field represented in Fourier series with $\underline{\gamma}_{\underline{k}} = \overline{\underline{\gamma}}_{-\underline{k}}$ (reality condition) and $\underline{k}\cdot\underline{\gamma}_{\underline{k}} = 0$ (incompressibility condition); here $\underline{k}$ has components that are integer multiples of the "lowest momentum" $k_0 = \frac{2\pi}{L}$. Then consider the equation:

$$\dot{\underline{\gamma}}_{\underline{k}} = -\vartheta(\underline{k})\underline{\gamma}_{\underline{k}} - i\sum_{\underline{k}_1+\underline{k}_2=\underline{k}}(\underline{\gamma}_{\underline{k}_1}\cdot\underline{k}_2)\,\Pi_{\underline{k}}\,\underline{\gamma}_{\underline{k}_2} + \underline{g}_{\underline{k}} \tag{2.8}$$

where the $\underline{k}$'s take only the values $0 < |\underline{k}| < K$ for some *momentum cut–off* $K > 0$ and $\Pi_{\underline{k}}$ is the projection on the plane orthogonal to $\underline{k}$. This is an equation that defines a "truncation on the momentum sphere with radius K of the equations (1.1)" if:

$$\begin{aligned}
&\vartheta(\underline{k}) = -\nu\,\underline{k}^2 \qquad \text{NS case}\\
&\vartheta(\underline{k}) = -\beta\,\underline{k}^2 \qquad \text{GNS case}
\end{aligned} \tag{2.9}$$

For simplicity we may suppose, in this no obstacles cases, that the mode $\underline{k} = \underline{0}$ is *absent*, *i.e.* $\underline{\gamma}_{\underline{0}} = \underline{0}$: this can be done if, as we suppose, the external force $\underline{g}$ does not have a zero mode component (*i.e.* if it has zero average).

In order that the resulting cut–off equations be physically acceptable, and supposing that $\underline{g}_{\underline{k}} \ne \underline{0}$ only for $|\underline{k}| \sim k_0$, one shall have to fix K large. For instance in the NS case it should be much larger than the *Kolmogorov scale* $K = (\eta\nu^{-2})^{1/4}$, where $\nu\eta$ is the average dissipation of the solutions to (2.6) with $K = +\infty$ (determined on the basis of heuristic dimensional considerations by $\eta \sim |\underline{g}|^2 L^2 \nu^{-2}$): see [11].

It is easy, in the no obstacles cases, to express the coefficients β for the cut off equations as $\beta \equiv \beta_e + \beta_i$:

$$\begin{aligned} \beta_e &= \frac{\sum_{\underline{k}\ne\underline{0}} \underline{k}^2 \underline{g}_{\underline{k}} \cdot \overline{\underline{\gamma}}_{\underline{k}}}{\sum_{\underline{k}} \underline{k}^4 |\underline{\gamma}_{\underline{k}}|^2} \\ \beta_i &= \frac{-i \sum_{\underline{k}_1+\underline{k}_2+\underline{k}_3=\underline{0}} \underline{k}_3^2 (\underline{\gamma}_{\underline{k}_1} \cdot \underline{k}_2)(\underline{\gamma}_{\underline{k}_2} \cdot \underline{\gamma}_{\underline{k}_3})}{\sum_{\underline{k}} \underline{k}^4 |\underline{\gamma}_{\underline{k}}|^2} \end{aligned} \tag{2.10}$$

where the $\underline{k}$'s take only the values $0 < |\underline{k}| < K$ for some *momentum cut–off* $K > 0$ and $\Pi_{\underline{k}}$ is the orthogonal projection on the plane perpendicular to $\underline{k}$.

The above two equations should be again *equivalent* in the natural sense that the stationary distribution $\mu_{\nu,ns}$ of the (truncated) NS equation with viscosity parameter ν and the stationary distribution $\mu_{\eta,gns}$ of the GNS equations should give the same averages to "most observables" if the value η is fixed to be the same as the average of the dissipation in NS equations, [10].

A key remark suggested by the above equivalence statements, which are in fact only conjectures, is that the stationary distributions can be regarded as statistical ensembles and the conjectures state that the same thermostatting mechanism can be equivalently modeled by completely different equations.

Furthermore the same thermostatting mechanism can be modeled by a "usual" irreversible equation of motion and *also* by a reversible one.

Much in the same way as in equilibrium statistical mechanics the same system can be described by the microcanononical or the canonical ensembles. Of course as in classical equilibrium mechanics the ensembles are not completely equivalent as one can always find quantities that have a completely different behavior in two "equivalent" systems (*e.g.* the energy fluctuations are quite different in the canonical and microcanonical ensembles even when they are equivalent).

The real problem is, however, *how to find the stationary distributions*?

This has been an outstanding problem in the last century: a solution for it was proposed in 1973 by Ruelle but it went unnoticed. Probably because it was formulated (later, *e.g.* in [12], and) in a way that a concrete application seemed far away.

Recently the Ruelle's principle has been shown to imply far reaching consequences. We therefore proceed to formulate it in a modern form having in mind to apply it (as an example) to the systems $(\mathcal{C}, S)$ obtained by timed observations on the above described family of systems.

In [13],[7],[14] the above systems, at least when *chaotic i.e.* when showing at least one positive Lyapunov exponent, are supposed to verify following hypothesis:

Chaotic hypothesis: A reversible many particle system in a stationary state can be regarded as a transitive Anosov system for the purpose of computing the macroscopic properties.

This means that the attractor verifies Axiom A and it is assumed to be smooth and

hyperbolic in the strict mathematical sense described in the kinematics section. At zero forcing we suppose, for compatibility with the ergodic hypothesis, the system to be a transitive hamiltonian Anosov system so that the attractor is the full phase space.

In the quoted references it is argued that the chaotic hypothesis should be considered in the same way as the *ergodic hypothesis* in equilibrium statistical mechanics. It is assumed as correct *even* in cases in which it cannot be mathematically strictly valid: but this can be done only for the purpose of deriving statistical properties of a few relevant observables. An analogue of this procedure is the derivation of the second law (*heat theorem*) from ergodicity (*i.e.* from the microcanonical ensemble) in equilibrium statistical mechanics: the law is derived by supposing ergodicity and it is assumed valid even when the ergodic hypothesis is obviously false (*e.g.* for the free gas in a box), [1].

It might be surprising that something can be concretely derived from this assumption: in fact it is very ambitious and far reaching and we shall see it has implications.

And it is clear that any consequence of such a general assumption must be a parameterless prediction, *i.e.* a *universal law*. Examples of this type of deductions are well known: the most remarkable is perhaps the mentioned *heat theorem* of Boltzmann which derives the second law of thermodynamics from the ergodic hypothesis (*i.e.* from the microcanonical ensemble).

In nonequilibrium thermodynamics there are no universally accepted laws that we could use as a reference and a test of the theory: except perhaps the Onsager reciprocity and the fluctuation dissipation theorems. The latter are derived in very many different ways and there is a general agreement about their correctness, see [15].

The Onsager reciprocal relations and the fluctuation dissipation relations (also called Green-Kubo formulae) have been subject of many unsuccessful, or only partially successful, attempts to extend them to non zero (or large) external forcing. In fact they "only" express relations between derivatives with respect to external fields *evaluated at* 0 *fields*!

We shall argue that the above chaotic hypothesis not only implies the Onsager relations (when they are valid, *i.e.* at zero fields) but that it also leads to an extension of them, in the form of a general parameterless statement, that will be called the *fluctuation theorem*.

§3 Reversible dissipation and the fluctuation theorem. Markov partitions (coarse graining and symbolic dynamics).

Note that the examples in §2 are pairs of dynamical systems which are *conjectured* to represent the statistical properties of the same system and the pairs have been deliberately selected so that each pair consists of a manifestly dissipative equation and of a manifestly reversible equation.

It is easy to check that the second equation of the example 1) is reversible: the operation $i:(\underline{p}\,,\,\underline{q}\,)\to(-\,\underline{p}\,,\,\underline{q}\,)$ induces a time reversal in phase space in the sense that the solution flow S_t verifies $iS_t=S_{-t}i$. This implies that the timing map S (on any time independent timing event) verifies $iS=S^{-1}i$.

Likewise the map $i:(\underline{p}\,,\,\underline{B}\,,\,\underline{q}\,,\,\underline{b}\,)\to(-\,\underline{p}\,,-\,\underline{B}\,,\,\underline{q}\,,\,\underline{b}\,)$ induces a time reversal symmetry in the phase space of the second equation in the second example.

And the second equation of the third example admits also the time reversal symmetry $i:\,\underline{u}\,\to-\,\underline{u}\,.$

In this section we shall restrict attention to such reversible systems and we shall suppose that i is an isometry (not restrictive, see §1) and that the chaotic hypothesis is verified.

Let x be a point on the phase space $\mathcal{C}$ of the systems that we study through timed observations and therefore are dynamical systems $(\mathcal{C},S)$ of the type considered in §1. We define the jacobian matrix $J(x)=\partial S(x)$ and regard it as a linear map of the tangent

plane T_x onto T_{Sx}. We also define the action of $\partial S(x)$ on the stable and unstable planes T^u_x, T^s_x as a linear map $J_u(x)$ or $J_s(x)$ onto T^u_{Sx}, T^s_{Sx}.

We call $\Lambda_u(x), \Lambda_s(x)$ the determinants of the jacobians $J_u(x), J_s(x)$: their product differs from the determinant $\Lambda(x)$ of $\partial S(x)$ by the ratio of the sine of the angle $a(x)$ between the planes T^s_x, T^u_x and the sine of the angle $a(Sx)$ between T^s_{Sx}, T^u_{Sx}. We set:

$$\Lambda_{u,\tau}(x) = \prod_{j=-\tau/2}^{\tau/2-1} \Lambda_u(S^j x), \quad \Lambda_{s,\tau}(x) = \prod_{j=-\tau/2}^{\tau/2-1} \Lambda_s(S^j x), \quad \Lambda_\tau(x) = \prod_{j=-\tau/2}^{\tau/2-1} \Lambda(S^j x) \tag{3.1}$$

hence $\Lambda(x) = \frac{\sin a(Sx)}{\sin a(x)} \Lambda_s(x)\Lambda_u(x)$.

The time reversal symmetry implies that $W^s_{ix} = iW^u_x, W^u_{ix} = iW^s_x$ and:

$$\begin{aligned} &\Lambda_\tau(x) = \Lambda_\tau(ix)^{-1}, \qquad \Lambda_{s,\tau}(ix) = \Lambda_{u,\tau}(x)^{-1}, \qquad \Lambda_{u,\tau}(ix) = \Lambda_{s,\tau}(x)^{-1} \\ &\sin a(x) = \sin a(ix) \end{aligned} \tag{3.2}$$

The quantity $\sigma(x) = -\log \Lambda(x)$ will be called the *entropy production* per timing event so that $e^{-\sigma(x)}$ is the *phase space volume contraction* per event.

It seems that $\sigma(x)$ has all the properties that one may wish for the extension to non equilibrium thermodynamics of the ordinary entropy, even though not everybody would agree on this statement.

For instance it is well established in many numerical experiments that the time average of $\sigma(x)$ is ≥ 0 and > 0 when the external fields are non zero. In zero field $\sigma(x)$ *vanishes* (an expression of Liouville's theorem).

Recently the general non negativity of the time average of $\sigma(x)$ has become a theorem (by Ruelle) that can be rightly called, I feel, the *H–theorem* of reversible non equilibrium statistical mechanics, [16]. More generally one can consider systems that are smooth only piecewise and extend the notion of SRB distribution so that, if the latter does not have nontrivial vanishing Lyapunov exponents, it can be also shown, *c.f.r* [16], that the average of $\sigma(x)$ can vanish on a SRB distribution[2] has a density with respect to the volume.

Therefore we shall call *dissipative* the systems for which the time average of $\sigma(x)$ is positive, [17].

We call $\sigma_\tau(x)$ the partial average of $\sigma(x)$ over the part of trajectory centered at x (in time): $S^{-\tau/2}x, \ldots, S^{\tau/2-1}x$. Then we can define the *dimensionless entropy production* $p = p(x)$ via:

$$\sigma_\tau(x) = \frac{1}{\tau} \sum_{j=-\tau/2}^{\tau/2-1} \sigma(S^j x) \stackrel{def}{=} \langle\sigma\rangle_+ p \tag{3.3}$$

where $\langle\sigma\rangle_+$ is the infinite time average $\int_{\mathcal{C}} \sigma(y)\mu_+(dy)$, if μ_+ is the forward statistics of the volume measure, and τ is any integer.

The *chaotic hypothesis* implies a *fluctuation theorem* which, [7], is a property of the fluctuations of the entropy production rate in dissipative systems; the (dimensionless) finite time average $p = p(x)$ has a statistical distribution $\pi_\tau(p)$ with respect to the stationary state distribution μ_+ such that the following limit exists:

[2] In general there may be many of them but, by the chaotic hypothesis assumed here, there is one and only one.

$$\lim_{\tau\to\infty}\frac{1}{\tau\langle\sigma\rangle_+}\log\pi_\tau(p)=-\zeta(p)\tag{3.4}$$

for all p's in the domain $(-p^*,p^*)$ where $p(x)$ can vary, and:

$$\frac{\zeta(p)-\zeta(-p)}{p\langle\sigma\rangle_+}=-1\tag{3.5}$$

provided the attractor is dense on phase space *and* the system is reversible. This is the *fluctuation theorem* of [13]: later we indicate its extension to cases in which the attractor is smaller than the whole phase space. Under our assumptions the function $\zeta(p)$ exists and is real analytic in $p\in(-p^*,p^*)$, [17].

The above relation holds in the whole domain of variability of p (which is in general a bounded variable because $\mathcal{C}$ is a bounded manifold): and therefore it is a surprising parameterless prediction.

The proof of the fluctuation theorem goes to the core of the structure of chaotic systems and is very enlightening. The key point is that Anosov (and Axiom A) systems have a dynamics that can be *easily* transformed into a symbolic dynamics, [2].

One can in fact find a partition of phase space $\mathcal{C}$ into *parallelograms.* A parallelogram will be a small set with a boundary consisting of pieces of the stable and unstable manifolds joined together as described below. The smallness has to be such that the parts of the manifolds involved look essentially "straight": *i.e.* the sizes of the sides have to be small compared to the smallest radii of curvature of the manifolds W^u_x and W^s_x, as x varies in $\mathcal{C}$.

Therefore let δ be a length scale small compared to the minimal (among all x) curvature radii of the stable and unstable manifolds. Let $W^{u,\delta}_x, W^{s,\delta}_x$ be the connected parts of W^u_x, W^s_x containing x and contained in a sphere of radius δ. Let us first define a *parallelogram* E in the phase space $\mathcal{C}$, to be denoted by $\Delta^u\times\Delta^s$, with center x and axes Δ^u, Δ^s with Δ^u and Δ^s small connected surface elements on W^u_x and W^s_x containing x. Then E is defined as follows.

Consider $\xi\in\Delta^u$ and $\eta\in\Delta^s$ and suppose that the intersection $\xi\times\eta\equiv W^{s,\delta}_\xi\cap W^{u,\delta}_\eta$ is a unique point (this will be so if δ is small enough and if Δ^u, Δ^s are small enough compared to δ as we can assume, because the stable and unstable manifolds are "smooth" and transversal).

The set $E=\Delta^u\times\Delta^s$ of all the points generated in this way when ξ,η vary arbitrarily in Δ^u,Δ^s will be called a parallelogram (or rectangle), if the boundaries $\partial\Delta^u,\partial\Delta^s$ of Δ^u and Δ^s as subsets of W^u_x and W^s_x, respectively, have zero surface area on the manifolds on which they lie. The sets $\partial_u E\equiv\Delta^u\times\partial\Delta^s$ and $\partial_s E=\partial\Delta^u\times\Delta^s$ will be called the *unstable* or *horizontal* and *stable* or *vertical* sides of the parallelogram E.

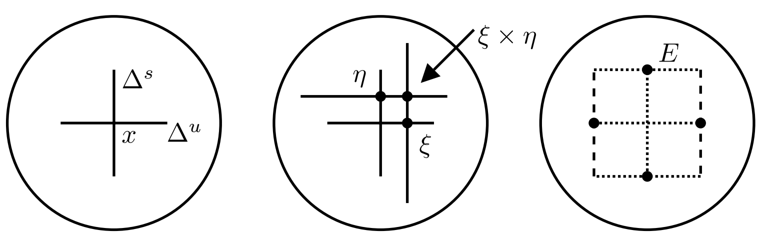

Fig. 1: The circles are a neighborhood of x of size very small compared to the curvature of the manifolds; the first picture shows the axes; the intermediate picture shows the $\times$ operation and $W^{u,\delta}_\eta, W^{s,\delta}_\xi$ (the horizontal

and vertical segments through η and ξ, respectively, have size δ); the third picture shows the rectangle E with the axes and the four marked points are the boundaries $\partial\Delta^u$ and $\partial\Delta^s$. The picture refers to a two dimensional case and the stable and unstable manifolds are drawn as flat, i.e. the Δ's is very small compared to the curvature of the manifolds. Travsversality of W^u_x, W^s_x is pictorially represented by drawing the surfaces at 90^o angles, from [7].

Consider now a partition $\mathcal{E} = (E_1, \ldots, E_{\mathcal{N}})$ of $\mathcal{C}$ into $\mathcal{N}$ rectangles E_j with pairwise disjoint interiors. We call $\partial_u\mathcal{E} \equiv \cup_j \partial_u E_j$ and $\partial_s\mathcal{E} \equiv \cup_j \partial_s E_j$: these are called respectively the *unstable boundary* of $\mathcal{E}$ and the *stable boundary* of $\mathcal{E}$, or also the horizontal and vertical boundaries of $\mathcal{E}$, respectively.

We say that $\mathcal{E}$ is a *Markov partition* if the transformation S acting on the stable boundary of $\mathcal{E}$ maps it into itself ($S\partial_s\mathcal{E} \subset \partial_s\mathcal{E}$) and if, likewise, the map S^{-1} acting on the unstable boundary maps it into itself ($S^{-1}\partial_u\mathcal{E} \subset \partial_u\mathcal{E}$).

The actual construction of the SRB distribution then proceeds from the important geometric result of the theory of Anosov systems expressed by what we shall call "Sinai's first theorem", [2]:

Theorem: every transitive Anosov system admits a Markov partition $\mathcal{E}$, as fine as wished.

The term "fine" refers to the maximum size $\delta_{\mathcal{E}}$ of the sets $E \in \mathcal{E}$ which can be fixed *a priori* to be smaller than any prefixed length ("as fine as wished").

The theorem can be extended to imply the existence of more special Markov partitions: for instance to show the existence of Markov partitions with any one of the following three properties (the last shows that the first two can be realized simultaneouly and will play a key role in our analysis):

(1) The construction of $\mathcal{E}$ can be done, [18], so that the horizontal axes of E_j *all lie on* W^u_O (and the vertical on W^s_O) and their union is a set that can be obtained from a single small connected surface element $\overline{\Delta}$ of W^u_O (resp. $\overline{\Delta}'$ of W^s_O) containing O by dilating it with a high iterate S^Q of the time evolution S. In other words the union $\cup_j \Delta^u_j$ of the horizontal axes of the parallelograms $E_j \in \mathcal{E}$ can be regarded as a single connected surface which in turn is a good finite representation of the attracting set.

Likewise the union of the stable axes can be regarded as a large connected part of the stable manifold W^s_O.

(2) If the reversibility property holds it is clear that $i\mathcal{E}$ is also a a Markov partition. This follows from the definition of Markov partition and from the fact that reversibility implies:

$$W^s_x = iW^u_{ix} \tag{3.6}$$

(3) The definition of a Markov partition also implies that the intersection of two Markov partitions is a Markov partition, hence it is clear that there are Markov partitions $\mathcal{E}$ that are reversible in the sense that $\mathcal{E} = i\mathcal{E}$ (if this is not true one can just intersect $\mathcal{E}$ and its i–image $i\mathcal{E}$).

Once a Markov partition $\mathcal{E} = (E_1, \ldots, E_{\mathcal{N}})$ is given one can associate with each point $x \in \mathcal{C}$ the *history* $\underline{h}\,(x) = (\ldots, h_{-1}, h_0, h_1, \ldots)$ of the motion of x on $\mathcal{E}$: it is the sequence of the names of the elements of $\mathcal{E}$ that x visits at time j: $S^j x \in E_{h_j}$.

This is well defined for all $x \in \mathcal{C}$ which do not visit in their evolution the boundaries of the elements of $\mathcal{E}$: the latter points form a set of zero volume and can be disregarded for the purposes of our analysis (having zero probability of being μ_0–randomly chosen). The expansion and contraction properties of the sides of the parallelograms imply that if the partition $\mathcal{E}$ is *fine enough* (precisely if the images under S of each $E \in \mathcal{E}$ intersects any other $E' \in \mathcal{E}$ in a connected set) they are small enough only one point can have a

given history and two points that have histories with identical symbols h_j for $|j| < N$ are close within a distance $O(e^{-\lambda N})$.

The symbolic representation is "as good" as the decimal representation of (the cartesian coordinates of) the points: provided we shall always suppose the Markov partition to be "fine enough" (in the above sense) it follows that the points $x \in \mathcal{C}$ are determined by the digits of their symbolic representation $\underline{h}$ with an exponential precision, in the sense that they get close at exponential rate in the number of agreeing digits. It is however much better than the usual decimal representations because it is *adapted* to the dynamics. The map S in fact just becomes the translation of the the history $\underline{h}(Sx) = \vartheta \underline{h}$ if ϑ is the left shift, obviously.

The above are properties that the history of x on $\mathcal{E}$ shares with most fine partitions of phase space. There is however a property that is enjoyed by the ordinary decimal representations of the cartesian coordinates and that is *a priori* missing in the symbolic representations that are derived from the histories on a generic partition of phase space.

Namely the ordinary representation of the cartesian coordinates by decimal digits has the remarkable property that not only each point can be represented symbolically but also *any sequence of digits* determines a point.

The latter property does not hold in general for the symbolic representation of a point via the history of its motion along a partition. In fact if we define the *compatibility matrix* $M_{hh'}$ for the partition $\mathcal{E}$ by setting $M_{hh'} = 1$ if $SE^0_h \cap E^0_{h'} \neq \emptyset$ (here E^0 denotes the interior of a set E) and $M_{hh'} = 0$ otherwise, then $\underline{h}$ can be the history of a point only if $M_{h_j h_{j+1}} \equiv 1$. We call such sequences *compatible* by nearest neighbours.

Note that if a system is a transitive Anosov system the matrix M has the *mixing property*, namely M^n has strictly positive matrix elements for all n large enough.

But it is clear that the compatibility condition is not the only one: one should define compatibility conditions linking three neighboring times $j, j+1, j+2$ or four and more. This shows that constructing a symbolic dynamics representation of a given dynamical system is a very non trivial problem. The understanding of the compatibility conditions essentially requires a complete solution of the dynamics in order to tell *a priori* which are the symbolic sequences that can actually represent a point, *i.e.* it requires the construction of a *code* of the phase space into a space of sequences.

What is special about Markov partitions is that we can *a priori* tell *which are the possible histories* of points in $\mathcal{C}$. They are simply *all* the sequences compatible by nearest neighbours: the space of such sequences will be called $\mathcal{K}_M$. This is *very remarkable* as in some sense it solves completely the problem of studying the motions of a transitive Anosov system.

Whatever the transitive Anosov system is the time evolution can be represented simply as a shift on the space $\mathcal{K}_M$ of biinfinite sequences of symbols subject only to nearest neighbour compatibility conditions ("subshift of finite type" with a suitable compatibility matrix). Therefore the Anosov systems play the same role that the harmonic oscillators play in perturbation theory: most systems are not Anosov systems, but since we "fully" understand Anosov systems they become the paradigm of non integrable motion just as the harmonic oscillators are the paradigm of integrable motions, [2].

§4 *Volume measure. Forward and backward SRB distributions.*

An immediate consequence of the symbolic dynamics representation of the points of $\mathcal{C}$ in the cases in which $(\mathcal{C}, S)$ is a transitive Anosov system is that the Liouville measure (*i.e.* the volume measure) on $\mathcal{C}$ and its statistics are easily expressible in symbolic terms.

We consider the partition $\mathcal{E}_T = \cap_{-T}^{T} S^{-j}\mathcal{E}$ obtained by intersecting the images under S^j,

$j = -T, \ldots, T$ of $\mathcal{E}$. Then $\mathcal{E}_T$ is still a Markov partition and it is time reversal invariant if $\mathcal{E}$ is.

We note that the parallelograms of $\mathcal{E}_T$ can be labeled by strings of symbols $h_{-T}, \ldots, h_T$ and they consist of the points x such that $S^k x \in E_{h_k}$ for $-T \le k \le T$. In other words the parallelograms consist of those points x which in their time evolution visit at time k the parallelogram h_k (one usually says that these are the points whose *symbolic dynamics* on $\mathcal{E}$ coincide, at the times ("sites") k between $-T$ and T).

In a parallelogram any point can be regarded as *center*. However there are special points that are usually taken as centers because they play a special role. Accordingly we shall take as center $x_{h_{-T},\ldots,h_T}$ of a parallelogram $E_{h_{-T},\ldots,h_T}$ a point whose symbolic dynamics string $\underline{j}$ at the times k with $|k| > T$ is fixed in a standard way; *i.e.* by defining h_k for $k > T$ (respectively $k < -T$) as a compatible sequence depending only on h_T (respectively h_{-T}). We cannot in general make the simple choice of continuing $h_{-T}, \ldots, h_T$ at times $k > T$ or $k < -T$ with a fixed symbol because this may lead to an incompatible sequence (hence the choice that we make is "the simplest" possible: it leaves some arbitrariness because there are many strings of symbols that start with a given symbol. The arbitrariness is however irrelevant for what follows).

The possibility of the above continuation of finite strings into infinite strings relies on the mixing property of the matrix M, consequence of the transitivity of $(\mathcal{C}, S)$.

Another simple choice of the center is to pick a periodic point in $E_{h_{-T},\ldots,h_T}$: this can be easily done by continuing the string $h_{-T}, \ldots, h_T$ beyond T in a standard way (in the above sense) to $h_{-T-n}, \ldots, h_{T+n}$ so that $M_{h_{T+n}h_{-T-n}} = 1$ and then continuing it *periodically*.

We can now define probability distributions on the phase space $\mathcal{C}$ simply by defining probability distributions on the space $\mathcal{K}_M$ of compatible sequences with compatibility matrix M of a fine Markov partition $\mathcal{E}$.

The simplest way of constructing probability distributions on $\mathcal{K}_M$ is to think of the symbolic sequences as *spin* chains on a one dimensional lattice (the time in our case) and to assign an energy to each chain and construct the corresponding Gibbs probability distribution.

We recall that an energy function for a chain of spins is a function $\ell_0(\underline{h})$, to be called the energy of interaction of the spin at 0 with the other spins of the chain, which has *short range*, *e.g.* for some $k, \kappa > 0$ is is $|\ell_0(\underline{h}) - \ell_0(\underline{h}')| \le k e^{-\kappa|N|}$ if $\underline{h}$ and $\underline{h}'$ agree on the sites j with $|j| < N$.

In other words an energy function has short range if it depends on a finite number of spins ("Ising chains") or if it depends *"exponentially little"* on the "far" spins.

Then one defines a Gibbs distribution with energy functions $\{\ell_j(\underline{h})\}$ by setting:

$$\int \tilde{\mu}(d\underline{h}')\tilde{F}(\underline{h}') = \lim_{T\to\infty} \frac{\sum_{h_{-T}\ldots,h_T} \tilde{F}(\underline{h}) e^{-\sum_{j=-T}^{T-1} \ell_j(S^j\underline{h})}}{\sum_{h_{-T}\ldots,h_T} e^{-\sum_{j=-T}^{T-1} \ell_j(S^j\underline{h})}} \equiv \lim_{T\to\infty} \tilde{\mu}_T(\tilde{F}) \tag{4.1}$$

here the infinite sequence $\underline{h}$ is a "standard" (in the above sense) continuation of the string $h_{-T}, \ldots, h_T$ to an infinite compatible sequence.

If ℓ_j verifies mild translation invariance properties, *e.g.* if $\ell_j(\underline{h}) = \ell_+(\underline{h})$ for all $j \ge 0$ and $\ell_j(\underline{h}) = \ell_-(\underline{h})$ for all $j < 0$, the limit exits and is unique.

The metric dynamical system $(\mathcal{K}_M, \vartheta, \tilde{\mu})$ has very strong cluster properties, for instance it is exponentially mixing in the sense that, if $\langle\cdot\rangle_\sim$ denotes $\tilde{\mu}$–average the correlation:

$$\langle\tilde F(S^j\cdot)\tilde F(\cdot)\rangle_\sim-\langle\tilde F(S^j\cdot)\rangle_\sim\langle\tilde F(\cdot)\rangle_\sim\xrightarrow[j\to\pm\infty]{}0 \tag{4.2}$$

and the convergence to zero is exponentially fast, for all functions on $\mathcal{K}_M$ with short range in the above sense.

One can construct in this way three gibbsian probability distributions that we call $\tilde\mu_0,\tilde\mu_+,\tilde\mu_-$. If $\underline{h}\longleftrightarrow x(\underline{h})$ is the code that associates with each compatible history $\underline{h}\in\mathcal{K}_M$ the point $x\in\mathcal{C}$ with that history, the distributions $\tilde\mu,\tilde\mu_\pm$ are the Gibbs distributions with respective potentials:

$$\begin{aligned}\ell_j(\underline{h})=&\begin{cases}\lambda_+(\underline{h})=\log\Lambda_u(x(\underline{h})) & \text{if } j>0\\ \lambda_-(\underline{h})=-\log\Lambda_s(x(\underline{h})) & \text{if } j<0\end{cases}\\ \ell_j(\underline{h})=&\lambda_+(\underline{h})\\ \ell_j(\underline{h})=&\lambda_-(\underline{h})\end{aligned}\tag{4.3}$$

Since the map $\underline{h}\to x(\underline{h})$ is such that $d(x(\underline{h}),x(\underline{h}'))<Be^{-\lambda N}$ if the histories $\underline{h},\underline{h}'$ agree between time $-N$ and time N, and since the functions $\Lambda_u(x),\Lambda_s(x)$ are Hölder continuous (see §1), it follows immediately that $\ell_j(\underline{h})$ have short memory and therefore they define Gibbs states $\tilde\mu_0,\tilde\mu_+,\tilde\mu_-$.

The above distributions $\tilde\mu$ can be coded back to distributions on $\mathcal{C}$ simply by using the code $\underline{h}\to x(\underline{h})$ between $\mathcal{K}_M$ and $\mathcal{C}$ and setting: $\mu(x(G))=\tilde\mu(G)$ for a generic set G.

The following theorem by Sinai is the fundamental result, [2]:

Theorem 2: Coding back into distributions on $\mathcal{C}$ the distributions $\tilde\mu_0,\tilde\mu_\pm$ one obtains three distributions μ_0,μ_+,μ_-. The first μ_0 is proportional to the volume, and in general it is not S–invariant, while the two others are the forward and backward statistics of the volume.

Thus this theorem solves completely the problem of the existence of the SRB distributions $\mu)\pm$. It also gives us a concrete and usable representation for the SRB distribution via (4.3) and (4.1).

The choice of the "standard" continuation of $h_{-T},\ldots,h_T$ to $\underline{h}\in K_M$ in (4.1) is quite arbitrary: but the above theorem implies that it is immaterial. In particular a very popular choice is the periodic continuation which leads to the alternative representation of the SRB distributions μ_+ and μ_- known as the *periodic orbit expansion.* Although usually (and inexplicably) surrounded by an aura of mystery this is a representation that is easier to communicate to unfavourable audiences as it does not require the symbolic dynamics analysis (*on which, however, it rests*):

$$\int F(x)\mu_+(dx)=\lim_{T\to\infty}\frac{\sum_{S^{2T+1}x=x}F(x)e^{-\sum_{j=-T}^{T-1}\lambda_+(S^jx)}}{\sum_{S^{2T+1}x=x}e^{-\sum_{j=-T}^{T-1}\lambda_+(S^jx)}}\tag{4.4}$$

which is a formula that is an immediate (non trivial) consequence (if one takes for granted the theory of one dimensional Gibbs states) of the previous (4.1),(4.3).

The periodic orbit representation of measures on on $\mathcal{C}$ has the disadvantage that the volume distribution μ_0 does not have a natural representation: only μ_+ (and μ_-) are easily represented. This is a drawback as (4.1),(4.3) make very transparent the relationship between $\mu_0,\mu_\pm$ and makes it obvious why $\mu_\pm$ are the forward and backward statistics of μ_0.

It is also possible to write (4.1) with $\ell_j=\lambda_+$ "directly" as a distribution on $\mathcal{C}$ as:

$$\int \mu_+(dx)F(x) = \lim_{T\to\infty} \mu_T(F) = \lim_{T\to\infty} \frac{\sum_j e^{-\sum_{k=-T}^{T-1} \lambda_+(x_j)(S^k x_j)} F(x_j)}{\sum_j e^{-\sum_{k=-T}^{T-1} \lambda_+(x_j)(S^k x_j)}} \tag{4.5}$$

where the sum runs over the elements E_j of the pavement $S^T\mathcal{E} \vee \ldots \vee S^{-T}E_{h_T}$, $\underline{h}$ is the extension of $(h_{-T},\ldots,h_T)$ to a string in $\mathcal{K}_M$ and $x_j \equiv x(\underline{h})$. The function F is any smooth (*i.e.* Hölder continuous) function on $\mathcal{C}$.

The relation (4.5) is also written as:

$$\int \mu_+(dx)F(x) = \lim_{T\to\infty} \frac{\sum_j \Lambda_{u,2T}(x_j)^{-1}F(x_j)}{\sum_j \Lambda_{u,2T}(x_j)^{-1}} \equiv \lim_{T\to\infty} \mu_T(F) \tag{4.6}$$

because by definition the jacobian determinant $\Lambda_{u,t}(x)$ of the jacobian of the map S^t as a map of $W^u_{S^{-t/2}x}$ to $W^u_{S^{t/2}x}$ is, see (3.1), $\exp\sum_{k=-t/2}^{t/2-1}\lambda_+(x_j)(S^k x_j)$.

We now show informally why (4.6) implies the fluctuation theorem, [13],[7],[17].

We first evaluate the probability, with respect to $\mu_{\tau/2}$ of (4.6), of $\sigma_\tau(x)/\langle\sigma\rangle_+ \in I_p$ divided by the probability (with respect to the same distribution) that $\sigma_\tau(x)/\langle\sigma\rangle_+ \in I_{-p}$. This is:

$$\frac{\sum_{j,\,a_\tau(x_j)=p} \overline{\Lambda}^{-1}_{u,\tau}(x_j)}{\sum_{j,\,a_\tau(x_j)=-p} \overline{\Lambda}^{-1}_{u,\tau}(x_j)} \tag{4.7}$$

Since $\mu_{\tau/2}$ in (4.6) is only an approximation to μ_+ *an error is involved in using* (4.7) as a formula for the same ratio computed by using the true μ_+ instead of $\mu_{\tau/2}$.

It can be shown that this error can be estimated to affect the result only by a factor bounded above and below uniformly in τ, p, [13]2. This is a remark technically based on the thermodynamic analogy pointed out in (4.1) above.

We now try to establish a one to one correspondence between the addends in the numerator of (4.7) and the ones in the denominator, aiming at showing that corresponding addends have a *constant ratio* which will, therefore, be the value of the ratio in (4.7).

This is possible because of the reversibility property: it will be used in the form of its consequences given by the relations (3.2).

The ratio (4.7) can therefore be written simply as:

$$\frac{\sum_{E_j,a_\tau(x_j)=p} \overline{\Lambda}^{-1}_{u,\tau}(x_j)}{\sum_{E_j,a_\tau(x_j)=-p} \overline{\Lambda}^{-1}_{u,\tau}(x_j)} \equiv \frac{\sum_{E_j,a_\tau(x_j)=p} \overline{\Lambda}^{-1}_{u,\tau}(x_j)}{\sum_{E_j,a_\tau(x_j)=p} \overline{\Lambda}_{s,\tau}(x_j)} \tag{4.8}$$

where $x_j \in E_j$ is the center in E_j. In deducing the second relation we make use of the existence of the time reversal symmetry i and of (3.5), and assume that the centers $x_j, x_{j'}$ of E_j and $E_{j'} = iE_j$ are chosen so that $x_{j'} = ix_j$.

It follows then that the ratios between corresponding terms in the ratio (4.8) is equal to $\overline{\Lambda}^{-1}_{u,\tau}(x)\overline{\Lambda}^{-1}_{s,\tau}(x)$. This differs from the reciprocal of the total change of phase space volume over the τ time steps between the point $S^{-\tau/2}x$ and $S^{\tau/2}x$ only because it does not take into account the ratio of the sines of the angles $a(S^{-\tau/2}x)$ and $a(S^{\tau/2}x)$ formed by the stable and unstable manifolds at the points $S^{-\tau/2}x$ and $S^{\tau/2}x$, see footnote [1]. But $\overline{\Lambda}^{-1}_{u,\tau}(x)\overline{\Lambda}^{-1}_{s,\tau}(x)$ will differ from the actual phase space contraction under the action of S^τ, as a map between $S^{-\tau/2}x$ and $S^{\tau/2}x$, by a factor that can be bounded between B^{-1} and B with $B = \max_{x,x'} \frac{|\sin a(x)|}{|\sin a(x')|}$ which is finite by the transversality of the stable and unstable manifolds.

Now for all the points x_j in (4.8), the reciprocal of the total phase space volume change over a time τt_0 is $e^{\sigma_\tau(x_j)\langle\sigma\rangle_+\tau}$, which (by the constraint imposed on the summation labels $\sigma_\tau/\langle\sigma\rangle_+ = p$) equals $e^{\tau\langle\sigma\rangle_+ p}$. Hence the ratio (4.7) will be $e^{\tau\langle\sigma\rangle_+ p}$. It is important to note that there are two errors ignored here, as pointed out in the discussion above. They imply that the argument of the exponential *is correct up to* p, τ *independent corrections* (which are in fact observed in the experiment as fig.3 of [19] shows, or as it is shown by Fig. 2 below). One should note that other errors may arise because of the approximate validity of our main chaotic assumption (which states that things go "as if" the system was Anosov): they may depend on N and we do not control them except for the fact that, if present, their relative value should tend to 0 as $N \to \infty$: there may be (and very likely there are) cases in which the chaotic hypotesis is not reasonable for small N (*e.g.* systmes like the Fermi-Pasta-Ulam chains) but it might be correct for large N. We also mention that for some systems with small N for which the chaotic hypothesis may be already regarded as valid (*e.g.* model 1 with $N = 1$ in [20]).

The p independence of the ratio (3.5) is therefore a key test of the theory (and it should hold with corrections of order $O(\tau^{-1})$ if $\zeta(p)$ is evaluated by its finite τ approximation $\zeta_\tau(p)$).

The fluctuation theorem is not immediately applicable in many cases. For instance in the cases in which the attractor is an axiom A attractor smaller than the whole phase space.

If however we imagine that the chaotic hypothesis holds in the stronger sense that the system verifies the Axiom C then we can use that the map $\tilde{\imath}$ that maps the backward attractor $\mathcal{A}_-$ into the forward attractor $\mathcal{A}_+$ and combine it with the time reversal symmetry i to generate a map $i^* = \tilde{\imath} i$ which maps the two poles $\mathcal{A}_+$ and $\mathcal{A}_-$ into themselves (see §2). The new map i^* is *a time reversal symmetry* that anticommutes with the time evolution, *i.e.* $Si^* = i^* S^{-1}$ *leaving* $\mathcal{A}_+$ *invariant.*

Thus the fluctuation therem holds for such systems because we can think that $(\mathcal{A}_+, S)$ is a reversible transitive Anosov system to which we can apply the above argunents.

Of course if we take the above viewpoint the phase space contraction that the theorem referes to should *no longer* be the contraction of the total volume but *just* the contraction of the surface measure of $\mathcal{A}_+$ which will now play the role of μ_0.

The latter result might be difficult to use as the attractor $\mathcal{A}_+$ is very likely to be an unreachable object. And in fact to be applied new properties must hold. It is remarkable that such poperties had already been observed in some special systems, much earlier than the fluctuation theorem. We discuss the applications of the theorem in §4.

Note that the theorem is a *large fluctuation* theorem: therefore it is very hard to test. The measure of how large are the fluctuations can be grasped also by the attempts that have been proposed in the literature to link the phenomena that occurr in models like the first one in §2 and the *Loschmidt paradox*, [21]. Clearly the deviations of p from the average are large deviations that, in order to be neasured, certainly involve observations of currents opposite to the field and violating the second law.

§*5. Applications: reversible conduction, fluids. Onsager reciprocity. Dynamical ensembles. Irreversible systems.*

The first applications are numerical tests. The basic result is [19] which was the origin of the whole story: the chaotic hypothesis was formulated to find a theoretical explanation of the results of this important experiment. The latter authomatically provides a first test of the hypothesis.

Experiments designed to test the hypthesis have since been performed. I quote here

the first of them, [14], that studies the reversible conduction model in example 1, §2. The following graphs are the results of tests of the fluctuation theorem in a 2 or 10 particles system of hard spheres enclosed in a periodic box containing two hard sphere fixed obstacles and subject to a *very* large electric field ($E = 1$ which is huge in physical units if one thinks of the particles as electrons in a crystal).

The following are the graphs of the empirically determined $x(p) = (\zeta(-p) - \zeta(p))/\langle\sigma\rangle_+$. The case of 2 particles and semiperiodic boundary conditions yields:

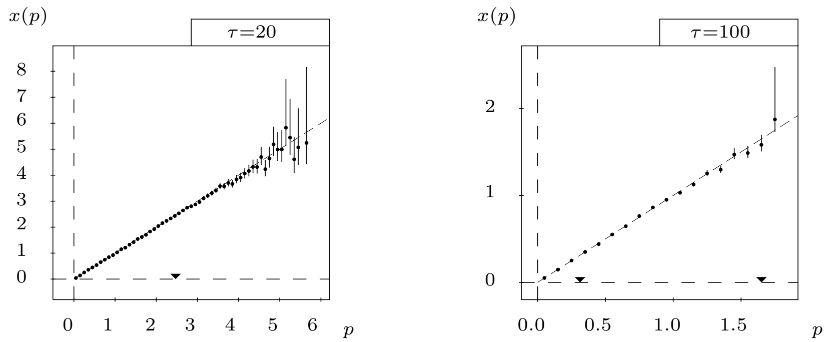

Fig.1: The graphs for the fluctuation theorem test $N = 2$; the dashed line is the fluctuation theorem prediction, $\tau = 20, 100$. The arrows mark the point at distance $\sqrt{\langle (p-1)^2 \rangle}$ from 1, from [14].

And the case of 10 particles and semiperiodic boundary conditions gives:

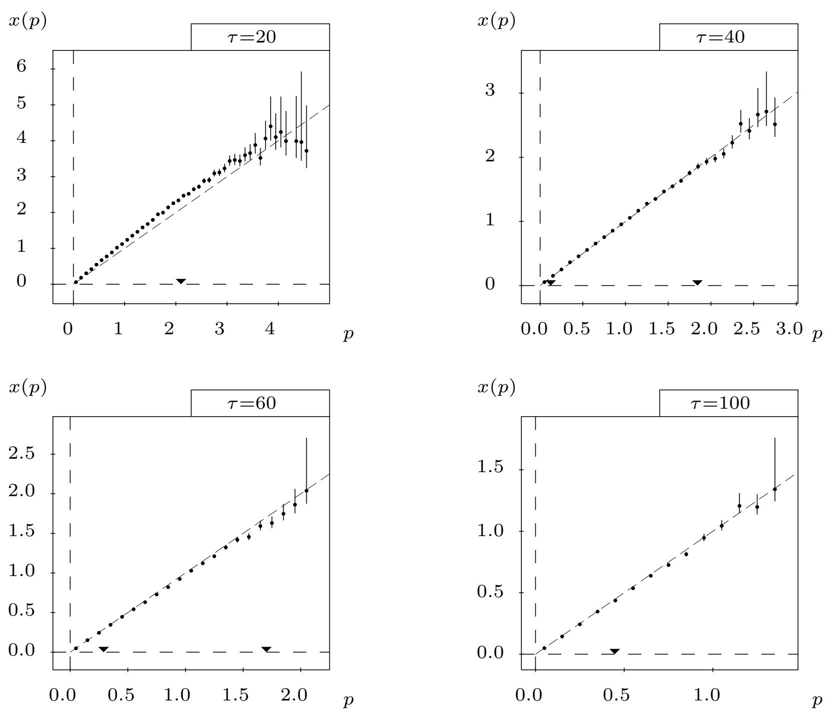

Fig.2: The linear fluctuation test, $\tau = 20, 40, 80, 100$ for $N = 10$. The dashed line is the fluctuation theorem prediction for $\tau = +\infty$. The arrows mark the point at distance $\sqrt{\langle (p-1)^2 \rangle}$ from 1, from [14].

According to the theorem the graphs should be a straight line with slope 1. The graphs show the data as well as the theoretical line with slope 1 (dashed). The function $\zeta(p)$, hence $x(p)$, in fact depends on τ, the time interval over which the fluctuations are ob-

served and should be more properly denoted $\zeta_\tau(p)$. We show the graphs of $x(p)$ versus p for various τ: the slope 1 is exact only for $\tau = +\infty$ but we see that it can already be observed for reasonable τ–values.

The experiments in [14] showed various other interesting phenomena: namely the fact that for large fields the attractor is certainly smaller than the full phase space.

The interpretation of the fact that in spite of that the experimental results still matched with (3.5), generated the idea of strengthening of the chaotic hypothesis by replacing the assumption that the system has attractors that can be regaded as Anosov systems by the stronger assumption that the system verifies Axiom C, which should have a general validity in time reversible systems: see [6].

If $(\mathcal{C}, S)$ is reversible and verifies Axiom C the map $\tilde{\imath}$ defined via geometric considerations in §2 will commute with i (by the covariance under i of the manifolds used to build it) and therefore the map $\tilde{\imath}i = i^*$ will leave invariant the poles $\Omega_\pm$ of the system and on them it will *anticommute* with S.

This means that even though time reversal symmetry is "lost" or "broken" on $\Omega_\pm$ there is *another* transformation (i^*) leaving $\Omega_\pm$ invariant and anticommuting with S, *i.e.* which plays the role of a time reversal symmetry for the restriction of S to Ω_+ (and Ω_-).

This implies that in reversible Axiom C systems *time reversal is undestructible*: if the attractor becomes smaller than the full phase space one can say that the time reversal symmetry is *spontaneously broken*. But on the attractor one can still define a map $i^* \neq i$ which anticommutes with time and leave the attractor invariant. This is the "real" time reversal for the attractor, or the *local time reversal*, see [6]: while the "true" time reversal appears as a broken symmetry to an observer that only observes motions that are attracted by Ω_+.

A situation reminiscent of the violation of time reversal in elemntary particles physics, where the time reversal symmetry, or T symmetry, is broken but there is another symmetry that chab=nges the sign to time, the TCP symmetry. Here i plays the role of T and i^* the role of TCP (and $\tilde{\imath}$ that of CP).

Hence the volume on the attractor will contract obeying the fluctutation theorem. This in itself is not sufficient to study the phase space volume fluctuations because, as said above, one does not have direct access to the surface area of the attractor (and to the attractor itself, as a matter of fact).

Then manifestly one would be back with an Anosov system (on a lower dimensional manifold) and a version of the fluctuation theorem would still hold. Furthermore one could say that this is only a different interpretation of the chaotic principle.

If this picture is correct we can write the phase space contraction rate (see (3.2) $\sigma(x) = \sigma_0(x) + \sigma_\perp(x)$ where $\sigma_0(x)$ is the contraction rate on the surface on which the attractor lies and $\sigma_\perp(x)$ is the contraction rate of the part of the stable manifold of the attractor which is not on the attractor itself (the angle between the part of the stable manifold sticking out of the attractor and the attractor itself is disregarded here as we think that it is bounded away from 0 and π since the attractor is compact).

Local time reversal will change the sign *only of* $\sigma_0(x)$ and the fluctuation theorem should apply to the fluctuations of σ_0. But $\sigma_0(x)$ is *not* directly accessible to measurement: nevertheless we can still study its fluctuations via the following *heuristic* analysis, [14].

Considerable help comes from a remarkable theorem that was discovered experimentally in [22], [23]. In the full phase space of the equations in the reversible example 1) in §1 the Lyapunov exponents verify a *pairing rule*. Namely if $2D = 4N - 2$ is the number of exponents and the first $D = 2N - 1$ exponents, $\lambda_1^+, \ldots, \lambda_{2N-1}^+$, are ordered in decreasing order and the next $2N - 1$, $\lambda_1^-, \ldots, \lambda_{2N-1}^-$, are ordered in increasing order then:

$$\frac{\lambda_j^+ + \lambda_j^-}{2} = const \qquad \text{for all } j = 1, \ldots, D \tag{5.1}$$

The constant will be called "*pairing level*" or "*pairing constant*": it must be $\frac{1}{2D}\langle\sigma\rangle_+$.

The pairing rule, in fact, formally holds in the present example as a consequence of [24].

In the cases in which (5.1) has been proved, [24], it holds also in a far *stronger* sense: the *local Lyapunov exponents*, of which the Lyapunov exponents are the averages, are paired as in (3.2) to a constant that is j independent but, of course, is dependent on the point in phase space. We call this the *strong pairing rule*. From the proof in [24] of the pairing rule one sees that the jacobian matrix $J = \partial S$ of the map S is such that $\sqrt{J^* J}$ has D pairs of eigenvalues ($D = 2N - 1$ in the case of example 1, §2) and the logarithms of each pair add up to $\sigma(x)$ (see also (2.4),(2.5)).

The simplest interpretation of this, consistent with the proposed attractor picture, is that pairs with elements of opposite signs describe expansion *on the manifold* on which the attractor lies. While the $M \le D$ pairs consisting of two negative exponents describe contraction of phase space *transversally to the manifold* on which the attractor lies.

Then we would have $\sigma_0(x) = (D - M)\sigma(x)$ and we should have a fluctuation law for the quantity $p(x)$ associated with $\sigma_0(x)$ defined by (2.6) and (2.7) with σ_0 replacing σ, *i.e.* (accepting the above heuristic argument, taken from [14]):

$$\sigma_{0\tau}(x) = \frac{(D - M)}{D}\langle\sigma\rangle_+ p(x) \tag{5.2}$$

i.e. a law identical to (2.10) *up to a correcting factor* $1 - \frac{M}{D}$:

$$\frac{1}{p\tau\langle\sigma\rangle_+} \log \frac{\pi_\tau(p)}{\pi_\tau(-p)} = (1 - \frac{M}{D})\, p \tag{5.3}$$

The graphs of Fig.2 are relative to an experiment in which we see that there may be one negative exponents in excess over the positive ones as said above.

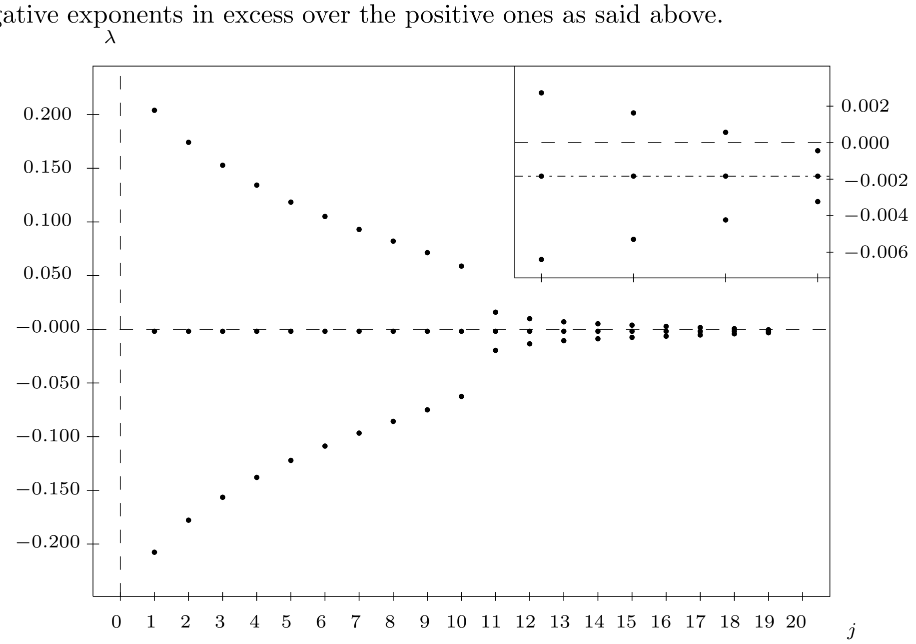

Fig.3 The 38 Lyapunov exponents, $N = 10$. Small picture magnifies the tail of the larger, showing better the pairing rule and that the 19–th exponent is slightly negative, from [14].

In fact a study of the exponents values in the case of 10 particles and semiperiodic bounday conditions gives the above diagram in which the Lyapunov exponents corresponding to $E = 1$ are drawn as a function of the field E. Which shows a very small exponent which is the higher member of a pair and, yet, is negative. The graphs of Fig. 2, however, show that the agreement of (4.1) with the experiment is within the errors: had there been no negative exponents we would have expected a slope 1. If there is one negative exponent in excess we expect a slope $1 - \frac{1}{19}$ which is within the error bars in Fig.2. An excess of 2 exponents would yield a slope of $1 - \frac{2}{19}$ which is *out* of the error bars.

Note that since the exponent smallest in modulus is so small we must expect that it yields a clear effect only after extremely long times have elapsed (*i.e.* for values of $\tau > 4.\,10^3$: totally out of computability). At larger fields the number of negative pairs increases considerably: but the flucututation theorem cannot be tested as the fluctuations become too improbabble to be measurable, even with the largest computers.

The reversible case of example 2) has not been studied. Nor the reversible case of model 3). Studies are being performed at least for the model in example 3).

More generally, however, *we do not expect the pairing rule to hold*. For instance this is clear in the case of the second example in §2: in that case the Lyapunov exponents relative to the thermostat are obviously paired to 0 while those of the conduction particles have a negative total sum (by the H–theorem of Ruelle, [16], quoted before (3.3)).

Nevertheless some kind of pairing might still occurr. In such cases one could envisage that (5.1) is replaced by a relation like:

$$\frac{\lambda_j^+ + \lambda_j^-}{2} = \langle\sigma\rangle_+ \frac{c_j}{2\overline{D}} \tag{5.4}$$

where $\langle\sigma\rangle$ is the μ_+ average of the phase space contraction per unit time; and c_j is some suitable function of j, while $\overline{D} \stackrel{def}{=} \sum_j c_j$.

We therefore *define* c_j by the (5.4) without attempting at determining them *a priori*. Then one can think that (5.4) holds in a "almost local" form *in the sense that on a rapid time scale (5.4) becomes true also for the local exponents*. This means that, *up to an error that tends to zero very quickly with the time* τ, the logarithms of the eigenvalues of the matrix $(J_\tau^T(x)J_\tau(x))^{1/2\tau}$, with $J_\tau(x)$ being the jacobian matrix for the evolution operator V_τ at x, divided by τ verify $\frac{1}{2}(\lambda_j^+ + \lambda_j^-) = c_j\beta_\tau(x)$ with $\beta_\tau(x)$ denoting the average $\frac{1}{2\overline{D}\tau}\sum_{j=-\tau/2}^{\tau/2-1} \beta(S^j x)\,dt$.

This property, together with the Axiom C assumption, will then suffice to extend, in a suitable form, the validity of the predictions of the fluctuation theorem based on the pairing rule (*i.e.* to cases in which the attractor is smaller than phase space).

We first remark that the really relevant feature of the pairing rule, as far as the above applications are concerned, is not the constancy of the pairing *but, rather, the fact that some kind of pairing takes place on a fast enough time scale*, see [10].

Then if a local time reversal exists on the attractor (*i.e.* if the geometric Axiom C is assumed as well for the dynamics the fluctuations of the observable σ will have a "free energy" (or a "generalized sum of Lyapunov exponents" to adhere to the terminology in [25], [26]) $\zeta(p)$, in the sense of (3.4), with an odd part $p\,\overline{P}\,\langle\sigma\rangle_+$, with $\overline{P}$ equal to $1 - \frac{\sum_- c_j}{\sum c_j}$ where the $\sum_-$ runs ver the values of the j's to which correspond two negative Lyapunov exponents:

$$\frac{\zeta(-p)-\zeta(p)}{\langle\sigma\rangle_+ p} = \overline{P} \stackrel{def}{=} 1 - \frac{\sum_- c_j}{\sum c_j} \tag{5.5}$$

This is a property whose validity can be conceivably tested in, real or numerical, experiments. At least the linearity in p of $-(\zeta(p)-\zeta(-p))/\langle\sigma\rangle_+$ with a slope ≤ 1 should be observable.

Another application is the already mentioned relation between the fluctuation theorem and the Onsager reciprocity.

We imagine a system even more general than the ones in §2. We assume that it is a *reversible dissipative* system with several forces $\underline{G} = (G_1, \ldots, G_s)$ acting upon the particles (or on the fluid in case of fluid mechanics models).

The $\underline{G}$ are parameters, with dimension not necessaribly being that of a force, measuring the strength of the various causes of nonequilibrium (*e.g.* they could be electromotive fields as abve, or temperature differences or other as in the references [27]). We follow here the analysis in [28] rather than the one in [27].

We suppose that the phase space contraction rate can be written in increasing powers of $\underline{G}$:

$$\sigma(x) = \sum_{i=1}^{s} G_i J_i^0(x) + O(G^2) \tag{5.6}$$

this simply means that we suppose that at zero forcing, $\underline{G} = \underline{0}$, there is no phase space contraction and, in fact, we always suppose that in such case the system is conservative (although this is not really necessary) and $\mu_+|_{\underline{G}=\underline{0}}$ is time reversal invariant.

The interpretation of $\sigma(x)$ as a microscopic density of entropy production allows us to establish in a unambiguous way the duality between *thermodynamic fluxes* or *currents* and *thermodynamic forces*.

We define the *current* associated with the *force* G_i by $J_i(x) = \partial_{G_i}\sigma(x)$ and the *transport coefficients* by setting $L_{ij} = \partial_{G_j}\langle J_i(x)\rangle_+|_{\underline{G}=0}$ and we study L_{ij}. We want to show that the above ideas suffice to prove Onsager's reciprocal relations *i.e.* $L_{ij} = L_{ji}$.

The analysis is based on the remark that the fluctuation theorem can be *extended* to give properties of the *joint* distribution of the average of σ, (5.6), and of the corresponding μ_+–average of $G_j\partial_{G_j}\sigma$.

In fact we shall define the *dimensionless* current associated with the force G_j by $q = q(x)$:

$$\frac{1}{\tau}\int_{-\tau/2}^{\tau/2} G_j\partial_{G_j}\sigma(S_t x)dt \stackrel{def}{=} G_j\langle\partial_{G_j}\sigma\rangle_+\, q \tag{5.7}$$

where the factor G_j is introduced here only to keep σ and $G_j\partial_{G_j}\sigma$ with the same dimensions.

Then if $\pi_\tau(p,q)$ is the joint probability of p,q the *same* proof of the fluctuation theorem discussed above yields also that, if we define:

$$\lim_{\tau\to\infty}\frac{1}{\tau}\log\pi_\tau(p,q) = -\zeta(p,q) \tag{5.8}$$

then:

$$\frac{\zeta(-p,-q)-\zeta(p,q)}{p\langle\sigma\rangle_+} = 1 \tag{5.9}$$

The proof in [17] shows that $\zeta(p,q)$ is analytic in the interior of the domain of variability of p,q at least if this domain is open. The latter condition means that p,q are *independent variables* for all $\underline{G}$ small enough.

This is the "normal" case: it is violated essentially only if the forces are not really different although they are given different names, *e.g.* one could have only one force G_1 acting on the system but one could regard it as two equal forces. In this case σ and $G_1\partial_{G_1}\sigma$ would coincide (at least to first order in $\underline{G}$) and the p,q would not be independent.

We can compute $\zeta(p,q)$ as usual in statistical mechanics: by considering first the transform $\lambda(\beta_1,\beta_2)$:

$$\lambda(\underline{\beta}) = \lim_{\tau\to\infty}\frac{1}{\tau}\log\int e^{\tau(\beta_1\,(p-1)\langle\sigma\rangle_+ + \beta_2\,(q-1)\langle G_j\partial_{G_j}\sigma\rangle_+)}\pi_\tau(p,q)dpdq \tag{5.10}$$

and then the Legendre trasform:

$$\zeta(p,q) = \max_{\beta_1,\beta_2}\big(\beta_1\,(p-1)\langle\sigma\rangle_+ + \beta_2\,(q-1)\langle G_j\partial_{G_j}\sigma\rangle_+ - \lambda(\underline{\beta})\big) \tag{5.11}$$

The function $\lambda(\underline{\beta})$, $\underline{\beta} = (\beta_1,\beta_2)$, is evaluated by the *cumulant expansion.* For the purpose of shortening the notation introduce $X(x)$ as:

$$X(x) = \beta_1\sum_{t=-\tau/2}^{\tau/2-1}(\sigma(S^t x) - \langle\sigma\rangle_+) + \beta_2\sum_{t=-\tau/2}^{\tau/2-1}(G_j\partial_j\sigma(S^t x) - \langle G_j\partial_j\sigma\rangle_+) \tag{5.12}$$

so that X has μ_+–average 0 and the function $\lambda(\underline{\beta})$ is simply written as:

$$e^{\tau\lambda(\underline{\beta})} = \int e^{X(x)}\mu_+(dx) = \langle e^X\rangle_+ \tag{5.13}$$

But $\langle X\rangle_+ = 0$ by the definition so that the cumulant expansion for $\int e^{X(x)}\mu_+(dx)$ yields, forgetting $O(G^3)$:

$$\int e^{X(x)}\mu_+(dx) = e^{\frac{1}{2!}\langle X^2\rangle_+} \tag{5.14}$$

and $\langle X^2\rangle_+$ is a quadratic form $C\,\underline{\beta},\underline{\beta})$ in $\underline{\beta}$ (because X is linear in $\underline{\beta}$) with coefficients given by the cumulants:

$$\begin{aligned}
C_{11} =& \sum_{-\infty}^{\infty}\big(\langle\sigma(S^t\cdot)\sigma(\cdot)\rangle_+ - \langle\sigma(\cdot)\rangle_+^2\big)\\
C_{22} =& \sum_{-\infty}^{\infty}\big(\langle G_j\partial_{G_j}\sigma(S^t)\cdot G_j\partial_{G_j}\sigma(\cdot)\rangle_+ - \langle G_j\partial_{G_j}\sigma(\cdot)\rangle_+^2\big)\\
C_{12} =& C_{21} = \sum_{-\infty}^{\infty}\big(\langle\sigma(S^t)\cdot G_j\partial_{G_j}\sigma(\cdot)\rangle_+ - \langle\sigma\rangle_+\,\langle G_j\partial_{G_j}\sigma(\cdot)\rangle_+\big)
\end{aligned} \tag{5.15}$$

The four cumulants form a symmetric matrix that will be called C.

The functions $\lambda(\underline{\beta})$ and $\zeta(p,q)$ are related via a Legendre transform:

$$\zeta(p,q) = \max_{\underline{\beta}}\big(\underline{\beta}\cdot\underline{w} - \lambda(\underline{\beta})\big) \tag{5.16}$$

where $\underline{w} = \begin{pmatrix} (p-1)\langle\sigma\rangle_+ \\ (q-1)\langle G_j\partial_j\sigma\rangle_+ \end{pmatrix}$.

So that the maximum condition yields the equations for the value of $\underline{\beta}$ where the expression in parenthesis in (5.16) reaches the maximum $\underline{\beta} = \underline{\beta}_{\max}$:

$$\begin{aligned}(p-1)\langle\sigma\rangle_+ =&\partial_{\beta_1}\lambda(\underline{\beta})|_{\underline{\beta}=\underline{\beta}_{\max}}\\ (q-1)\langle G_j\partial_j\sigma\rangle_+ =&\partial_{\beta_2}\lambda(\underline{\beta})|_{\underline{\beta}=\underline{\beta}_{\max}}\end{aligned} \tag{5.17}$$

so that, always performing the computations by neglecting quantities of $O(G^3)$ (*i.e.* by using $\lambda(\underline{\beta}) = \frac{1}{2}(C\underline{\beta}, \underline{\beta})$):

$$\underline{w} = \begin{pmatrix} (p-1)\langle\sigma\rangle_+ \\ (q-1)\langle G_j\partial_j\sigma\rangle_+ \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \qquad \text{or} \quad \underline{\beta}_{\max} = C^{-1}\underline{w} \tag{5.18}$$

and from (5.16), evaluating the r.h.s. at the maximum point $\underline{\beta}_{\max}$, we get:

$$\zeta(p,q) = \underline{w}\cdot\underline{\beta}_{\max} - \lambda(\underline{\beta}_{\max}) = \frac{1}{2}(\underline{w}, C^{-1}\underline{w}) \tag{5.19}$$

Since:

$$C^{-1} = \frac{1}{\det C}\begin{pmatrix} C_{22} & -C_{12} \\ -C_{21} & C_{11} \end{pmatrix} \tag{5.20}$$

we see that (recall that C is symmetric):

$$\begin{aligned}&\zeta(p,q) = \frac{1}{2}(C^{-1})_{11}(p-1)^2\langle\sigma\rangle_+^2 + \\ &+\frac{1}{2}(C^{-1})_{22}(q-1)^2\langle G_j\partial_j s\rangle_+^2 + (C^{-1})_{12}(q-1)(p-1)\sigma_+\,G_j\partial_j\sigma_+\end{aligned} \tag{5.21}$$

and the odd terms in (p,q) have coefficients:

$$\begin{aligned}q \quad &\to \quad (-2(C^{-1})_{22}G_j\langle\partial_{G_j}\sigma\rangle_+ - 2(C^{-1})_{12}\langle\sigma\rangle_+)\,G_j\langle\partial_{G_j}\sigma\rangle_+ = 0\\ p \quad &\to \quad (-2(C^{-1})_{11}\langle\sigma\rangle_+ - 2(C^{-1})_{12}G_j\langle\partial_{G_j}\sigma\rangle_+)\langle\sigma\rangle_+) = -\langle\sigma\rangle_+\end{aligned} \tag{5.22}$$

where the r.h.s. *arise by imposing compatibility with the fluctuation theorem relation* (5.9).

The above two relations, after some simple algebra, are seen to imply:

$$\begin{aligned}\langle\sigma\rangle_+ =&\frac{1}{2}C_{11} + O(G^3)\\ \langle G_j\partial_{G_j}\sigma\rangle_+ =&\frac{1}{2}C_{12} + O(G^3)\end{aligned} \tag{5.23}$$

Then expanding both sides of (5.23) *to lowest order* (*i.e.* lowest non trivial, the second in this case) in the G_i's we get the Onsager relation and the Green Kobo formula.

In fact first we look at the r.h.s. of the first relation in (5.23): in C_{11} one simply replaces σ, see (5.15), by its expansion to first order, (5.6), and the r.h.s. becomes a quadratic form in $\underline{G}$ with coefficients given by:

$$\frac{1}{2}\int_{-\infty}^{\infty} dt\,\big(\langle J_i^0(S_t\cdot)J_j^0(\cdot)\rangle_+ - \langle J_i^0\rangle_+\langle J_j^0\rangle_+\big)\big|_{G=0} \tag{5.24}$$

On the other hand the expansion of $\langle\sigma\rangle_+$ in the l.h.s. of the first of (5.23) to second order in $\underline{G}$ gives:

$$\langle\sigma\rangle_+ = \sum G_i \partial_{G_i} \langle\sigma\rangle_+\big|_0 + \frac{1}{2}\sum_{ij} G_i G_j \big(\partial_{G_i}(\partial_{G_j}\langle\sigma\rangle_+)\big|_0 \tag{5.25}$$

the first term vanishes (by time reversal, or just because it is clear from the r.h.s. of (5.23) that $\langle\sigma\rangle_+$ is of second order in $\underline{G}$). The second term is the sum of $\frac{1}{2}G_iG_j$ times:

$$\begin{aligned}\partial_{G_i}\partial_{G_j}\int\sigma(x)\mu_+(dx) = \int\big(\partial_{G_i}\partial_{G_j}\sigma(x)\mu_+(dx) +\\ + \partial_{G_i}\sigma(x)\partial_{G_j}\mu_+(dx) + \partial_{G_j}\sigma(x)\partial_{G_i}\mu_+(dx) + \sigma(x)\partial_{G_i}\partial_{G_j}\mu_+(dx)\big)\end{aligned} \tag{5.26}$$

evaluated at $G = 0$; and the first term in the r.h.s. vanishes (by time reversal it changes sign, while μ_+ is invariant) the second and third terms are $\partial_j < J_i^0 >_+ |_0 + (i \longleftrightarrow j)$ and the fourth vanishes (because $\sigma = 0$ at $G = 0$): note that $(\partial_{G_j} < J_i^0 >_+)|_0$ is equal to $(\partial_{G_j} < J_i >_+)|_0$).[3]

Therefore by equating the r.h.s and the l.h.s.of (5.23), we get from the first of (5.23) the above expression (5.24) for the matrix $\frac{L_{ij}+L_{ji}}{2}$ giving Green–Kubo's formula for $i = j$ (but not the Onsager reciprocity nor the general Green–Kubo formula which would say that L_{ij} equals (5.24)).

The same type of analysis on the second of (5.23), *which, unlike the first relation, is not symmetric in the* $\underline{G}$ *'s as* j *is privileged*, leads to the "asymmetric relation":

$$L_{ji} \equiv \partial_{G_j}\langle\partial_{G_i}\sigma\rangle|_{\underline{G}=\underline{0}} = \frac{1}{2}\sum_{t=-\infty}^{\infty}\big(\langle J_i^0(S_t\cdot)J_j^0(\cdot)\rangle_+ - \langle J_i^0\rangle_+\langle J_j^0\rangle_+\big)\big|_{G=0} \tag{5.27}$$

which gives the general Green-Kubo formula, hence Onsager reciprocity and the fluctuation dissipation theorem.

Thus the Onsager relations are a consequence of the fluctuation theorem (not surprisingly) and of its (obvious) extension, (5.9), in the limit $G \to 0$, when combined with the cumulant expansion for entropy fluctuations. Those theorems and the fast decay of the $\sigma\sigma$ correlations are all natural consequences of the chatic hypothesis in reversible statistical mechanical or fluid mechanical systems.

Therefore while the Onsager reciprocity and Green–Kubo formulae (or fluctuation dissipation theorem) only hold around equilibrium, *i.e.* they are properties of G–derivatives evaluated at $G = 0$, and the cumulant expansion for $\lambda(\underline{\beta})$ is a general consequence of the correlation decay in Anosov systems, the fluctuation theorem also holds far from equilibrium, *i.e.* for large $\underline{G}$, and can be considered a generalization of the Onsager relations and of the Green–Kubo formulae, [28].

Evidence for the above interpretation of the fluctuation theorem arose in [14] in an effort to interpret the results of various numerical experiments and an apparent incompatibility

[3] This is again the same argument:

$$\begin{aligned}&\partial_j < J_i >_+ |_0 \equiv \partial_j(\int \partial_i\sigma(x)\mu_+(dx))|_0 = (\int \partial_j\partial_i\sigma(x)\mu_+(dx))_0 + (\int \partial_i\sigma(x)\partial_j\mu_+(dx))_0 =\\ &= 0 + (\int \partial_i\sigma(x)|_0\partial_j\mu_+(dx)|_0 = \partial_j < J_i^0 >_+)_0\end{aligned}$$

where the last step uses that J_i^0 being evaluated at $G = 0$ does not depend on G.

of the *a priori* known non gaussian nature of the distribution $\pi_\tau(p)$ and the "gaussian looking" empirical distributions.

The next question is what we can say about the irreversible models? The key is the already mentioned conjectured equivalence of the *corresponding* SRB distributions.

The equivalence conjectures seem to raise the possibility of a genral theory of non equilibrium statistical ensembles and of their equivalence.

The idea of using reversible equations to study rreversible behaviour is quite appealing as we have seen that there are by now quite a few results that can be obtained for reversible dissipative evolutions. Hopefully the theory will grow and we shall learn to use the chaotic hypothesis and the ensuing characterization of the SRB distributions. After all in equilibrium statistical mechanics it took a long time to develop the theory from the heat theorem to the universal phenomena at criticality.

Unfortunately the equivalence of the ensembles does not allow us to make predictions based on the fluctuation theorem about the fluctuations of the entropy production in irreversible systems, even when equivalent to reversible ones. This is so because the entropy production is a "global" observable that in the irreversible models is ususally fixed *a priori* to a given value or, at least, to a value related to a quantity fixed *a priori*.

Thus the fact that we cannot translate knowledge of the entropy fluctuations in reversible models to irreversible equivalent models becomes analogous to the impossibility of translating information on the energy fluctuations in the canonical enseble to informations (of any interest) on the energy fluctuations in the microcanonical ensemble.

Thus although equivalence can be tested, we have, so far, no special predictions to mention aside the obvious one that one should get the same result when measuring corresponding quantities (a rather non trivial and interesting property): the fluctuation theorem is not directly usable beyond the fact that it implies the Onsgaer reciprocity.

In [10] some applications have been proposed for the analysis of the models in the example 3) of §2 and they may be tested, perhaps, experimentally. But this is beyond the scopes of the present review.

The equivalence conjecture *on the other hand* gives us an opportunity to test how good the approximations of irreversible equations for macroscopic phenomena are.

In fact if a measurement (a hypothtical one as so far no measurements of this type seem to have been performed) of the entropy fluctuations agrees with the fluctuation theorem then this would mean that the reversible models are better models (compared to the irreversible ones) of the system as they give the same predictions for normal observables but for the entropy production fluctuations they give different predictions (compared to the ones of the irreversible models).

Appendix A1. An example of Axiom C system.

We give here an example, taken from [6], in which i^*, the local time reversal, arising in the applications can be easily constructed. The example illustrates what we think is a typical situation. The poles $\Omega_\pm$, Ω_-, will be two compact regular surfaces, identical in the sense that they will be mapped into each other by the time reversal i defined below.

If x is a point in $M_* = \Omega_+$ the generic point of the phase space will be determined by a pair (x, z) where $x \in M_*$ and z is a set of transversal coordinates that tell us how far we are from the attractor. The coordinate z takes two well defined values on Ω_+ and Ω_- that we can denote z_+ and z_- respectively.

The coordinate x identifies a point on the compact manifold M_* on which a reversible transitive Anosov map S_* acts (see [18]). And the map S on phase space is defined by:

$$S(x,z) = (S_* x, \tilde S z) \tag{A1.1}$$

where $\tilde S$ is a map acting on the z coordinate (marking a point on a compact manifold) which is an evolution leading from an unstable fixed point z_- to a stable fixed point z_+. For instance z could consist of a pair of coordinates v, w with $v^2 + w^2 = 1$ (*i.e.* z is a point on a circle) and an evolution of v, w could be governed by the equation $\dot v = -\alpha v,\ \dot w = E - \alpha w$ with $\alpha = Ew$. If we set $\tilde S z$ to be the time 1 evolution (under the latter differential equations) of $z = (v, w)$ we see that such evolution sends $v \to 0$ and $w \to \pm 1$ as $t \to \pm\infty$ and the latter are non marginal fixed points for $\tilde S$.

Thus if we set $S(x,z) = (S_* x, \tilde S z)$ we see that our system is hyperbolic on the basic sets $\Omega_\pm = M_* \times \{z_\pm\}$ and the future pole Ω_+ is the set of points (x, z_+) with $x \in M_*$; while the past pole Ω_- is the set of points (x, z_-) with $x \in M_*$.

Clearly the two poles are mapped into each other by the map $i(x, z_\pm) = (i^* x, z_\mp)$. But on each attractor a "local time reversal" acts: namely the map $i^*(x, z_\pm) = (i^* x, z_\pm)$.

The system is "chaotic" as it has an Axiom A attracting set with closure consisting of the points having the form (x, z_+) for the motion towards the future and a different Axiom A attracting set with closure consisting of the points having the form (x, z_-) for the motion towards the past. In fact the dynamical systems (Ω_+, S) and (Ω_-, S) obtained by restricting S to the future or past attracting sets are Anosov systems because $\Omega_\pm$ are regular manifolds.

We may think that in the reversible cases the situation is always the above: namely there is an "irrelevant" set of coordinates z that describes the departure from the future and past attractors. The future and past attractors are copies (via the global time reversal i) of each other and on each of them is defined a map i^* which inverts the time arrow, *leaving the attractor invariant*: such map will be naturally called the *local time reversal*.

In the above case the map i^* and the coordinates (x, z) are "obvious". The problem is to see that they are defined quite generally under the only assumption that the system is reversible and has a unique future and a unique past attractos that verify the Axiom A. This is a problem that is naturally solved in general when the system verifies the Axiom C of §1.

In the following section we shall describe the interpretation of i^* in terms of symbolic dynamics when the system verifies Axiom C: as one may expect the construction is simple but it is deeply related to the properties of hyperbolic systems such as their Markov partitions.

Appendix A2. Ising model analogy.

The fluctuation theorem as expressed by (2.10) and the subsequent comments on the Gaussian nature of the function $\pi_\tau(p)$ may seem somewhat strange and unfamiliar.

It is therefore worth pointing out that the phenomenon of a "linear fluctuation law" on the odd part of the distribution, in the sense of (4.1), without a globally Gaussian distribution, is in fact well known in statistical mechanics and probability theory. And an example, proposed in [14], of what the fluctuation theorem means in a concrete case, in which $\pi_\tau(p)$ is *not* Gaussian, can be made by using the Ising model on a 1 dimensional lattice Z.

We consider the space $\mathcal{C}$ of the *spin configurations* $\underline\sigma = \{\sigma_\xi\}$, $\xi \in Z$ and the map S that translates each configuration to the right (say). The "time reversal" is the map $i : \{\,\underline\sigma\,\} \to \{-\underline\sigma\,\}$ that changes the sign to each spin.

The probability distribution that approximates the SRB distribution is the finite volume Gibbs distribution:

$$\mu_\Lambda(\underline{\sigma}) = \frac{\exp\left(J\sum_{j=-T}^{T}\sigma_j\sigma_{j+1} \,+\, h\sum_{j=-T}^{T}\sigma_j\right)}{normalization} \tag{A2.1}$$

where $\Lambda = [-T, T]$ is a large interval, $J, h > 0$. The configuration $\underline{\sigma}$ outside Λ is distributed independently on the one inside the box Λ, to fix the ideas.

Calling $\langle m\rangle_+$ the average magnetization in the thermodynamic limit we define the magnetization in a box $[-\frac{\tau}{2}, \frac{\tau}{2}]$ to be $M_\tau = \tau\langle m\rangle_+ p$ and we look at the probability distribution $\pi_\tau^T(p)$ of p in the limit $T \to \infty$. The Gibbs distribution corresponding to the limit of (A2.1) will play the role of the SRB distribution. Calling this limit probability $\pi_\tau(p)$ it is easy to see that:

$$\frac{\pi_\tau(p)}{\pi_\tau(-p)} \xrightarrow[\tau\to\infty]{} e^{2\tau h\langle m\rangle_+ p} \tag{A2.2}$$

This is in fact obvious if we take the two limits $T \to \infty$ and $\tau \to \infty$ *simultaneously* by setting $T = \frac{\tau}{2}$. In such a case, if $\sum_{\underline{\sigma}\,;\,p}$ denotes summation over all the configurations with given magnetization in $[-T, T]$, *i.e.* such that $\sum_{j=-\frac{\tau}{2}}^{\frac{\tau}{2}-1}\sigma_j = \langle m\rangle_+ p$ the distribution (A2.1) gives us immediately that:

$$\frac{\pi_\tau^T(p)}{\pi_\tau^T(-p)} = \frac{\sum_{\underline{\sigma}\,,\,p}\exp J\sum_{j=-T}^{T}\sigma_j\sigma_{j+1} \,+\, h\sum_{j=-T}^{T}\sigma_j}{\sum_{\underline{\sigma}\,,\,-p}\exp J\sum_{j=-T}^{T}\sigma_j\sigma_{j+1} \,+\, h\sum_{j=-T}^{T}\sigma_j} \equiv e^{2\tau h\langle m\rangle_+ p} \tag{A2.3}$$

if we use the symmetry of the pair interaction part of the energy under the "time reversal" (*i.e.* under spin reversal).

The error involved, in the above argument, in taking $T = \frac{\tau}{2}$ rather then first $T \to \infty$ and then $\tau \to \infty$, can be easily corrected since the corrections are "boundary terms", and in one dimensional short range spin systems there are no phase transitions and the boundary terms have no influence in the infinite volume limit (*i.e.* they manifest themselves as corrections that vanish, as $T \to \infty$ followed by $\tau \to \infty$).

One may not like that the operation i commutes with S rather than transforming it into S^{-1}. Another example in which the operation i does also invert the sign of time is obtained by defining i as $\{i\,\underline{\sigma}\,\}|_j = -\sigma_{-j}$: the (A2.3) can be derived also by using this new symmetry operation.

The above examples show why there is *a priori* independence between any Gaussian property of $\pi_\tau(p)$ and the fluctuation theorem. The theory of the fluctuation theorem in [13] is in fact *based* on the possibility (discovered in [2]) of representing a chaotic system as a one dimensional short range system of interacting spins (in general higher that $\frac{1}{2}$); and the argument is, actually, very close to the above one for the Ising model with, however, a rather different time reversal operation. See [17] for mathematical details on the boundary condition question.

Acknowledgements: I are indebted to F. Bonetto, E.G.D. Cohen, P. Garrido, G . Gentile, G. Paladin for stimulating discussions and comments: the results reviewed here can be originally found in joint works or originated from common discussions. This work is an expanded version of a series of talks at the meeting *Let's face chaos through nonlinear dynamics*, at the University of Maribor, Slovenia, 24 june 1996– 5 july 1996. It is part of the research program of the European Network on: "Stability and Universality in Classical Mechanics", # ERBCHRXCT940460, and it has been partially supported also by CNR-GNFM and Rutgers University.

References.

[1] Gallavotti, G.: *Ergodicity, ensembles, irreversibility in Boltzmann and beyond*, Journal of Statistical Physics, **78**, 1571–1589, 1995.

[2] Sinai, Y.G.: *Gibbs measures in ergodic theory*, Russian Mathematical Surveys, **27**, 21–69, 1972. Also: *Introduction to ergodic theory*, Princeton U. Press, Princeton, 1977.

[3] Smale, S.: *Differentiable dynamical systems*, Bullettin of the American Mathematical Society, **73** , 747–818, 1967.

[4] Arnold, V., Avez, A.: *Ergodic problems of classical mechanics*, Benjamin, 1966.

[5] Ruelle, D.: *Elements of differentiable dynamics and bifurcation theory*, Academic Press, 1989.

[6] Bonetto, F., Gallavotti, G.: *Reversibility, coarse graining and the chaoticity principle*, mp_arc@ math. utexas. edu, #96–50.

[7] Gallavotti, G., Cohen, E.G.D.: *Dynamical ensembles in stationary states*, Journal of Statistical Physics, **78**, ,1995.

[8] Gallavotti, G.: *Equivalence of dynamical ensembles and Navier Stokes equations*, mp_arc 96-131; or chao-dyn 9604006, in print on Physics Letters A.

[9] Nelson, E.: *Dynamical theories of brownian motion*, Mathematical Notes, Princeton U. Press, 1967.

[10] Gallavotti, G.: *Dynamical ensembles equivalence in statistical mechanics*, in mp_arc@ math. utexas. edu #96-182, chao-dyn@xyz.lanl.gov #9605006, in print on Physica D.

[11] Landau, L., Lifshitz, V.: *Mécanique des fluides*, MIR, Moscow, 1966.

[12] Ruelle, D.: *A measure associated with Axiom A attractors*, American Journal of Mathematics, **98**, 619–654, 1976.

[13] Gallavotti, G., Cohen, E.G.D.: *Dynamical ensembles in nonequilibrium statistical mechanics*, Physical Review Letters, **74**, 2694–2697, 1995.

[14] Bonetto, F., Gallavotti, G., Garrido, P.: *Chaotic principle: an experimental test*, mp_arc @math. utexas. edu, # 96-154.

[15] de Groot, S., Mazur, P.: *Non equilibrium thermodynamics*, Dover, 1984.

[16] Ruelle, D.: *Positivity of entropy production in the presence of a random thermostat*, in mp_arc@ math. utexas. edu, # 96-167. See also *Positivity of entropy production in nonequilibrium statistical mechanics*, in mp_arc@ math. utexas. edu, # 96-166.

[17] Gallavotti, G.: *Reversible Anosov diffeomorphisms and large deviations.*, Mathematical Physics Electronic Journal, **1**, (1), 1995, (http:// www. ma. utexas. edu/ MPEJ/ mpej.htlm).

[18] Gallavotti, G.: *Topics in chaotic dynamics*, Lectures at the Granada school 1994, Ed. P. Garrido, J. Marro, in Lecture Notes in Physics, Springer Verlag, **448**, p. 271–311, 1995.

[19] Evans, D.J.,Cohen, E.G.D., Morriss, G.P.: *Probability of second law violations in shearing steady flows*, Physical Review Letters, **71**, 2401–2404, 1993.

[20] Chernov, N. I., Eyink, G. L., Lebowitz, J.L., Sinai, Y.: *Steady state electric conductivity in the periodic Lorentz gas*, Communications in Mathematical Physics, **154**, 569–601, 1993.

[21] Holian, B.L., Hoover, W.G., Posch. H.A.: *Resolution of Loschmidt's paradox: the origin of irreversible behavior in reversible atomistic dynamics*, Physical Review Letters, **59**, 10–13, 1987.

[22] Evans, D.J., Morriss, G.P.: *Statistical Mechanics of Nonequilibrium fluids*, Academic Press, New York, 1990.

[23] Evans, D.J.,Cohen, E.G.D., Morriss, G.P.: *Viscosity of a simple fluid from its maximal Lyapunov exponents*, Physical Review, **42A**, 5990–5997, 1990.

[24] Dettman, C.P., Morriss, G.P.: *Proof of conjugate pairing for an isokinetic thermostat*, N.S. Wales U., Sydney 2052, preprint, 1995.

[25] Frisch, U., Parisi, G.: in *Turbulence and predictability*, p. 84–87, ed. M. Ghil, R. Benzi, G.Parisi, Noth Holland, 1984.

[26] Bohr, T., Jensen M.H., Paladin, G., Vulpiani, A.: *Dynamical systems approach to turbulence*, Cambridge Nonlinear series, 1996.

[27] Gallavotti, G.:*Chaotic hypothesis: Onsager reciprocity and fluctuation–dissipation theorem*, J. Statistical Physics, **84**, 899–926, 1996. In print in Journal of Statistical Physics, 1996. See also: *Chaotic principle: some applications to developed turbulence*, archived in *mp_arc@math.utexas.edu*, #95-232, and *chao-dyn@ xyz.lanl.gov*, #9505013, in print in J. of Statistical Physics.

[28] Gallavotti, G.: *Extension of Onsager's reciprocity to large fields and the chaotic hypothesis*, mp_ arc 96-109; or chao-dyn 9603003, in print in Physical Review Letters.

Internet access: All the Author's quoted preprints can be found and freely downloaded (latest postscript version including corrections of misprints and errors) at:

`http://chimera.roma1.infn.it`

in the Mathematical Physics Preprints page.

e-mail address of author: giovanni@ipparco.roma1.infn.it